

Causation: Some Examples

- BILLIARDS A moving billiard ball strikes a stationary billiard ball (A), causing it to move (B).
 COGS In a chain of cogs, the rotation of the next-to-last cog (A) causes the rotation of the last cog (B).
 BOMB A hydrogen bomb detonates in Rochester at noon (A), destroying Ithaca two minutes later (B).
 MIRROR A photon in outer space strikes a mirror (A), causing it to reflect at the angle of incidence (B).

(Assume that $t_A < t_B$, that $\neg A \Box \rightarrow \neg B$, and that A *influences* B).

The Causal Placement Problem

- ELIMINATIVISM There are no causal facts (*yes there are*).
 PRIMITIVISM Causal facts do not supervene on fundamental physical facts (*yes they do*).
 PHYSICALISM Causal facts supervene on fundamental physical facts (*which ones?*).

The placement problem is non-trivial:

- LANGUAGE Causal locutions are absent from the language of fundamental physics.
 INSTANTANEITY The fundamental laws of physics describe *instantaneous* functional relationships.
 GLOBALITY The fundamental laws of physics express constraints on *global* states of physical systems.
 ASYMMETRY The fundamental laws of physics do not discriminate between past and future.

Two physicalisms:

- REDUCTIONISM Causal facts supervene on non-agential physical facts (*popular*).
 RELATIVISM Causal facts supervene on physical facts that include agential physical facts (*true*).

Anti-Reductionism

- 1 Either REDUCTIONISM is true or RELATIVISM is true.
- 2 If REDUCTIONISM is true then there exist non-agential physical facts able to ground the asymmetry of causation.
- 3 There do not exist non-agential physical facts able to ground the asymmetry of causation.
- 4 Therefore RELATIVISM is true (1, 2, 3, 4).

Some terminology: A law is *time reversal invariant* just in case for any allowed sequence of states $S_1, S_2 \dots S_{n-1}, S_n$ the sequence $S_n^T, S_{n-1}^T \dots S_2^T, S_1^T$ is also allowed, where S_i^T represents the time reversal transformation of S_i (e.g. BOMB^T). A law is *temporally oriented* if it determines physical state evolutions in only one temporal direction.

- Q1 In virtue of what does A cause B and not *vice versa*?
 Q2 In which if any of BILLIARDS^T, COGS^T, BOMB^T and MIRROR^T does B cause A?

- REDUCTIONISM₁ (A1) The direction of time. (A2) All of them.
Problems: No backwards causation by *fiat*; spooky epistemology; metaphysical profligacy.
 REDUCTIONISM₂ (A1) Macroscopic asymmetry. (A2) Please don't ask that question.
Problems: No micro-causation; no causal supervenience; macroscopic counterexamples.
 RELATIVISM (A1) Agential asymmetry. (A2) It depends on which agents we relativise to.
Problems: So far all we have is a Morgenbesser proof that P: If not P, what? Q maybe?

Pro-Relativism

- 1 The temporal asymmetry of causation is grounded in the temporal asymmetry of counterfactual dependence.
- 2 The temporal asymmetry of counterfactual dependence is partly grounded in agential physical facts.
- 3 If (1) and (2), then the temporal asymmetry of causation is partly grounded in agential physical facts.
- 4 Therefore RELATIVISM is true (1, 2, 3).

First Premise

Independent attractiveness of counterfactual theories of causation. *Objection:* The description of BILLIARDS justified belief that A caused B without justifying belief that $\neg A \Box \rightarrow \neg B$. *Reply:* The description justified belief in counterfactuals entailing that A is a *direct cause* of B:

DIRECT CAUSE X is a *direct cause* of Y with respect to some variable set V *iff* there exists a possible intervention on X that would change Y with all other variables in V besides X and Y held fixed at some value by interventions.

Second Premise

SEMANTICS $\neg A \Box \rightarrow \neg B$ *iff* some $\neg A$ -world where $\neg B$ holds is closer to the actual world than is any $\neg A$ -world where $\neg B$ does not hold.

DIVERGENCE The closest $\neg A$ -world is identical to the actual world until a time t_Δ before t_A , then diverges in such a way to make $\neg A$ true, then evolves according to the actual laws. Call this world the *divergence world* and the period of time between t_Δ and t_A the *transition period*.

Q1 In virtue of what is the divergence world the closest possible world?

Q2 Why isn't it the case that, for some event Δ in the transition period, $\neg A \Box \rightarrow \neg \Delta$?

LEWIS (A1) Asymmetry of overdetermination. (A2) Indeterminate due to ties for closeness.

Problems: There is no such asymmetry; there are not always ties.

LOEWER (A1) Asymmetry of statistical mechanics. (A2) It *is* the case, but all Δ are microscopic.

Problems: There is no such asymmetry; not all Δ are microscopic.

BENNETT (A1) Conceptual analysis tells me so. (A2) Indeterminate due to vagueness in the closeness relation.

Problems: We're doing metaphysics; vagueness in closeness isn't always reflected in Δ -indeterminacy.

RELATIVISM (A1) It reflects the structure of the interventions it is physically possible for agents like us to perform.

(A2) Δ is *part* of the hypothetical intervention, and so not the *result* of it.

Earman¹:

To build a world w' that is like w up to time t but different thereafter, I can simply imagine myself letting w "run on" up to time t (no sweat so far), creating a small localized miracle at t , and then resting from my brief labors while Nature takes Her course. But to build a world w'' that is like w after t but which diverges from w for time before t , I have to imagine myself changing records, traces, memories, etc., in w , and this is a sweaty job. Needless to say, however, this imagined labor is not a relevant commodity.

Yes it is!

¹John Earman. 1976. "Causation: A Matter of Life and Death", in *The Journal of Philosophy*, Vol. 73, No. 1, January 1976, pp. 5-25