Problems of a theoretical analysis of causation

Michael Baumgartner

University of Geneva

28 April, 2016

Overview

1. Problems of modern theories of causation
   - Goal of the philosophy of causation
   - Relata of causation
   - General vs. singular causation
   - Relational properties of the causal relation(s)
   - Realism vs. anti-realism
   - Causal principles
   - Pre-theoretical clarification of the analysandum
2. Candidate theories
   - Regularity theory
   - Counterfactual theory
   - Probabilistic theory
   - Transfer theory
   - Interventionist theory
3. Summary

Goal of the philosophy of causation

- A philosophical analysis of the causal relation searches for necessary and sufficient conditions for an occurrence/dependency to be of causal nature. What are the truth conditions of sentences as „x causes y“? The goal is to come up with a biconditional of the following form

$$x \text{ causes } y \text{ if, and only if, } P. \qquad (\Phi)$$

- *Reductionist* theories search for a P that is free of causal connotations, *non-reductionist* theories allow for a 'causally loaded' P.
- **Method:** candidate theories of type (Φ) are confronted with standardly accepted pre-theoretic causal judgements.
- **Problem:** pre-theoretic causal judgements are often ambiguous and not always consistent.

Examples

Walter goes on vacation. His neighbor agrees to water Walter's plants, but repeatedly forgets to do so. When Walter returns two weeks later, his plants are dead. Which of the following factors are causes of the plants' death?

- the neighbor's failure to water regularly

Examples

Walter goes on vacation. His neighbor agrees to water Walter's plants, but repeatedly forgets to do so. When Walter returns two weeks later, his plants are dead. Which of the following factors are causes of the plants' death?

- the neighbor's failure to water regularly
- insufficient water supply

Examples

Walter goes on vacation. His neighbor agrees to water Walter's plants, but repeatedly forgets to do so. When Walter returns two weeks later, his plants are dead. Which of the following factors are causes of the plants' death?

- the neighbor's failure to water regularly
- insufficient water supply
- the Pope's failure to water regularly

Examples

Walter goes on vacation. His neighbor agrees to water Walter's plants, but repeatedly forgets to do so. When Walter returns two weeks later, his plants are dead. Which of the following factors are causes of the plants' death?

- the neighbor's failure to water regularly
- insufficient water supply
- the Pope's failure to water regularly
- Walter's failure to give the plants to his (reliable) mother prior to departure

Examples

Walter goes on vacation. His neighbor agrees to water Walter's plants, but repeatedly forgets to do so. When Walter returns two weeks later, his plants are dead. Which of the following factors are causes of the plants' death?

- the neighbor's failure to water regularly
- insufficient water supply
- the Pope's failure to water regularly
- Walter's failure to give the plants to his (reliable) mother prior to departure
- Walter's purchase of the plants at Migros

Examples

Walter goes on vacation. His neighbor agrees to water Walter's plants, but repeatedly forgets to do so. When Walter returns two weeks later, his plants are dead. Which of the following factors are causes of the plants' death?

- the neighbor's failure to water regularly
- insufficient water supply
- the Pope's failure to water regularly
- Walter's failure to give the plants to his (reliable) mother prior to departure
- Walter's purchase of the plants at Migros
- Walter's birth

Examples

Walter goes on vacation. His neighbor agrees to water Walter's plants, but repeatedly forgets to do so. When Walter returns two weeks later, his plants are dead. Which of the following factors are causes of the plants' death?

- the neighbor's failure to water regularly
- insufficient water supply
- the Pope's failure to water regularly
- Walter's failure to give the plants to his (reliable) mother prior to departure
- Walter's purchase of the plants at Migros
- Walter's birth
- Big Bang

Examples

A delinquent is brought before a firing squad composed of two shooters. Both of them simultaneously hit him in the heart. The delinquent dies. (Each shot would have been lethal by itself.) Which of the following factors are causes of the delinquent's death?

- the shot of the first shooter
- the shot of the second shooter

Examples

Walter has many enemies and one bodyguard. His enemies have repeatedly tried to kill Walter by poisoning his morning coffee. The bodyguard has found out about the enemies' plans and has obtained an antidote, which he adds to Walter's coffee every morning. The antidote neutralizes the poison (if there is poison in the coffee in the first place) and itself has no side-effects. On one particular morning there is no poison in Walter's coffee. The bodyguard adds the antidote. Walter survives. Which of the following factors are causes of Walter's survival?

- the absence of poison in the coffee
- the presence of the antidote in the coffee

Goal of a theory of causation

- A theoretical analysis of causation cannot attempt to capture *all* pre-theoretic causal judgments.

- The goal must be to account for a maximally large consistent proper subset **K** of all pre-theoretic causal judgments. That is, a successful theory of causation provides a (Φ) that is satisfied if, and only if, the analysandum "x causes y" is replaced by sentences expressing judgments in **K**.

What are the relata of causation?

- **Widespread intuition 1:** causes and effects are occurrences, events, states of affairs in time and space.
- For example, the accident is the cause of Walter's injury.
- **Widespread intuition 2:** absences and omissions can often be causally interpreted as well, even though they are no events in time and space.
- For example, the absence of oxygen in Walter's blood is a cause of his death.

→ There are both *event* and *fact* theories of causation.

What are the relata of causation?

- **Widespread intuition 1:** causes and effects are occurrences, events, states of affairs in time and space.
- For example, the accident is the cause of Walter's injury.

- **Widespread intuition 2:** absences and omissions can often be causally interpreted as well, even though they are no events in time and space.
- For example, the absence of oxygen in Walter's blood is a cause of his death.

→ There are both *event* and *fact* theories of causation.

General vs. singular causation

- Two causal relations must be kept apart: one on type level, *general causation*, and one on token level, *singular causation*.
- The relation of general causation connects types (of events or facts), the relation of singular causation connects tokens (of events or facts).
  - "Smoking causes lung cancer" vs. "Walter's smoking causes Walter's lung cancer".
  - "An increase in prices causes a loss of purchasing power" vs. "The increase in prices in Switzerland in August 2009 causes a loss of purchasing power of 0.3% in Switzerland in August of 2009".
  - "Collisions with icebergs cause shipwrecks" vs. "The Titanic's collision with the iceberg causes the Titanic's shipwreck".

Relational properties of the causal relation(s)

- It is uncontroversial that both general and singular causation are *not symmetric*: $\neg\forall x\forall y(Cxy \rightarrow Cyx)$. [Singular causation is moreover often considered to be *asymmetric*: $\forall x\forall y(Cxy \rightarrow \neg Cyx)$].

- Furthermore: singular causation is clearly *irreflexive*: $\forall x\neg Cxx$; and general causation is *non-reflexive*: $\neg\forall xCxx$.

- With respect to transitivity of general and singular causation there are ambiguous pre-theoretic intuitions:

  Are general and singular causation *transitive* or not:
  $\forall x\forall y\forall z(Cxy \wedge Cyz \rightarrow Cxz)$?

Transitivity?

Walter goes hiking. On a steep slope he sees that a huge boulder has dislodged 50 meters above him and is rolling towards him. At the last moment, he manages to duck. The boulder just misses his head. Walter remains uninjured. The dislodging of the boulder is a cause of Walter's ducking, which in turn is a cause of his physical integrity. Is the dislodging of the boulder also a cause of Walter's physical integrity? (Hitchcock)

Transitivity?

Billy and Suzy are friends. Suzy is malicious, Billy is benevolent. Billy observes that Suzy has picked up a rock and is now taking a run-up to throw the rock into a nearby window. Billy runs towards his friend to keep her from throwing the rock, but he trips over a garden hose and falls. Suzy is completely oblivious of Billy, throws the rock, and shatters the window. Billy's tripping a cause of Suzy's throwing of the rock, which in turn is a cause of the shattered window. Is Billy's tripping a cause of the shattered window? (Hall)

Causal realism vs. causal anti-realism

Causal realism: Not only the relata of the causal relation exist, but also the relation itself. Causes and effects are related by some *de re* necessity; they a are connected by an existing 'causal bond' (Armstrong, Tooley).

Causal anti-realism: Only the relata of causation and their behavior exist. There is no *de re* necessity relating causes and effects; there are no 'causal bonds' (Hume, Kant).

Causal principles

Principle of determinism: Whenever the same types of causes occur, the same types of effects occur as well (same causes, same effects).

Principle of causality: Events/facts do not occur/are not the case without at least one of their causes (every event has a cause).

The conjunction of the principle of determinism and the principle of causality implies:

Causal determinism: Every state of the universe is causally determined by its preceding state and causally determines its successor state.

What are the properties of the notion(s) of causation to be analyzed?

- Does causation relate events or facts or both?
- Is causation transitive or not?
- Does the causal relation itself exist (in addition to its relata)?
- Is causation deterministic?
- Are there uncaused occurrences?

→ The inconsistent and often ambiguous pre-theoretic usage of causal notions must be cleared up before a philosophical analysis of these notions becomes possible (or fruitful).

→ Theorizing about causation requires finding a balance between pre-theoretic intuitions and theoretical virtues as consistency and unambiguity.

Regularity theory (presuppositions)

(a) The causal relation itself is not part of our ontology (**anti-realism**).

(b) **General causation** is the primary analysandum.

(c) **Universal regularities** between types of events are the primary analysans.

(d) Causation is deterministic.

- Moreover: Some regularity theories render causation transitive, others do not.

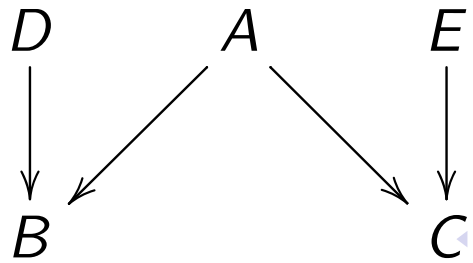
Regularity theory (Hume, Mill)

(I)

A is causally relevant to B iff A is a part of a sufficient condition AX_1 of B, which, in turn, is part of necessary condition (in disjunctive normal form) of B: $AX_1 \vee X_2 \vee \ldots \vee X_n \leftrightarrow B$.

(I) faces the following difficulties:

- Material conditionals contain redundancies, e.g. they are monotonic. That is, if AX_1 is sufficient for B, AX_1Z is likewise sufficient for B, where Z can be any arbitrary factor.
- (I) cannot distinguish between parallel effects of a common cause and genuine causal dependencies:



Regularity theory (modern)

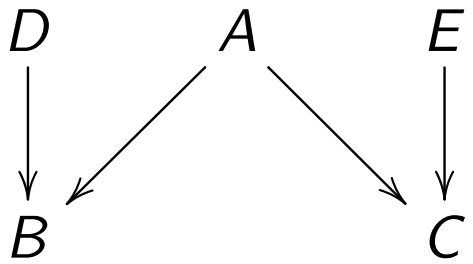
(II)

A is causally relevant to B iff:

(i) A is part of a minimally sufficient condition AX_1 of B,

(ii) AX_1 is contained in a disjunction $AX_1 \vee X_2 \vee \ldots \vee X_n$, $n \geq 2$, featuring other minimally sufficient conditions of B such that $AX_1 \vee X_2 \vee \ldots \vee X_n$ is minimally necessary for B.

- (II) eliminates all redundancies from regularities and can distinguish between parallel effects of a common cause and genuine causal dependencies:



Regularity theory (modern)

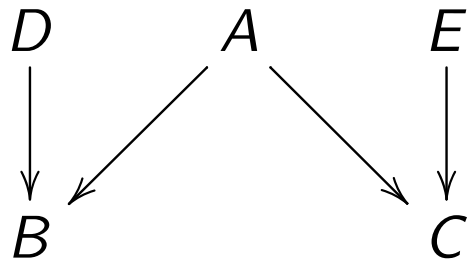
(II)

A is causally relevant to B iff:

(i) A is part of a minimally sufficient condition AX_1 of B,

(ii) AX_1 is contained in a disjunction $AX_1 \vee X_2 \vee \ldots \vee X_n$, $n \geq 2$, featuring other minimally sufficient conditions of B such that $AX_1 \vee X_2 \vee \ldots \vee X_n$ is minimally necessary for B.

- (II) eliminates all redundancies from regularities and can distinguish between parallel effects of a common cause and genuine causal dependencies:



$B\overline{D} \vee A \vee E$ is not minimally necessary for C.

Regularity theory (modern)

(II) faces the following difficulties:

- Universal regularities, as required by (II), are rather rare. Empirical data that are causally analyzed in the sciences hardly ever feature universal regularities.
- (II) presupposes the validity of the principle of determinism. That all causes determine their effects, however, is put into question by the standard interpretation of Quantum Mechanics.

Counterfactual theory (presuppositions)

(a) The causal relation itself is not part of our ontology (**anti-realism**).

(b) **Singular causation** is the primary analysandum.

(c) **Counterfactual dependencies** between token events are the primary analysans.

(d) Causation is deterministic.

- Moreover: Some counterfactual theories render causation transitive, others do not.

Counterfactual theory (conceptual preliminaries)

Counterfactual dependence: In a situation, in which events *a* and *b* occur, *b* is counterfactually dependent on *a* iff: had *a* not occurred, *b* would not have occurred either.

Counterfactual theory (conceptual preliminaries)

Counterfactual dependence: In a situation, in which events a and b occur, b is counterfactually dependent on a iff: had a not occurred, b would not have occurred either.

Truth condition of a counterfactual conditional: In a situation, in which events a and b occur, the counterfactual conditional "had a not occurred, b would not have occurred either" $(\neg a \; \Box\!\!\rightarrow \neg b)$ is true iff the non-a-world which is most similar to the actual world is also a non-b-world.

Counterfactual theory (Lewis)

(III)

a causes *b* iff

(i) *a* and *b* occur; and

(ii) there is a finite chain of (occurring) events $\langle a, x_1, \ldots, x_n, b \rangle$ such that each element of the chain counterfactually depends on its predecessor in the chain.

(III) faces the following difficulties:

- (III) fails in cases of *overdetermination* and cases of so-called *pre-emption*.
- The notion of counterfactual dependence is not more basic than the notion of causal dependence.
- The notion of counterfactual dependence is highly context-dependent.

Probabilistic theory (presuppositions)

(a) The causal relation itself is not part of our ontology (**anti-realism**).

(b) **General causation** is the primary analysandum.

(c) **Probabilistic dependencies** between types of events are the primary analysans.

(d) Causation is not deterministic.

(e) Causation is not transitive.

Probabilistic theory (conceptual preliminaries)

Prima facie cause: A at t' ($A_{t'}$) is a prima facie cause of B_t iff:

$$t' < t, \tag{1}$$

$$\mathrm{P}(A_{t'}) > 0, \tag{2}$$

$$\mathrm{P}(B_t \mid A_{t'}) > \mathrm{P}(B_t). \tag{3}$$

Screen off: A is screened off from B by C iff A does not make difference to the probability of B conditional on C, or formally:

$$\mathrm{P}(B \mid A \wedge C) = \mathrm{P}(B \mid C). \tag{4}$$

Spurious cause: $A_{t'}$ is a spurious cause of B_t iff $A_{t'}$ is a prima facie cause of B_t and there exists a third factor $C_{t''}$ such that:

$$t'' < t', \tag{5}$$

$$\mathrm{P}(B_t \mid A_{t'} \wedge C_{t''}) = \mathrm{P}(B_t \mid C_{t''}). \tag{6}$$

Probabilistic theory (Suppes)

(IV)

A is causally relevant to *B* iff *A* is a prima facie cause but not a spurious cause of *B*, i.e. iff *A* is a probability raiser of *B* and there does not exist a factor (instantiated before *A*) that screens off *A* from *B*.

(IV) faces the following difficulties:

- Not all causes are *probability raisers* of their effects.
- An understanding of probability in terms of relative frequency renders (IV) vulnerable to paradoxical frequency distributions (e.g. Simpson's Paradox).
- While in cases of *reducibly* indeterministic processes, common causes indeed always screen off their parallel effects, in cases of *irreducibly* indeterministic processes this does not hold generally.

Simpson's Paradox

	hair loss	no hair loss	total	rate
1 *male*				
Alopezin	7	3	10	70%
¬ Alopezin	18	12	30	60%
	25	15	40	
2 *female*				
Alopezin	9	21	30	30%
¬ Alopezin	2	8	10	20%
	11	29	40	
3 *male* $\vee$ *female*				
Alopezin	16	24	40	40%
¬Alopezin	20	20	40	50%
	36	44	80	

Transfer theory (presuppositions)

(a) The causal relation itself is part of our ontology (**realism**).

(b) **Singular causation** is the primary analysandum.

(c) **Transfer processes** between events are the primary analysans.

- Moreover: Transfer theories are non-committal with respect to whether causation is deterministic and transitive or not.

Transfer theory (Dowe)

(V)

a causes *b* iff

(i) *a* and *b* occur; and

(ii) there is a transfer of energy or momentum (or of some other conserved quantity) from *a* to *b*.

(V) faces the following difficulties:

- (V) cannot make sense of causation by absence.
- The whole area of social, political or economic processes cannot be causally interpreted against the background of a transfer theory.
- The notion of a transfer process is 'causally loaded', (V) hence is a non-reductive account.

Interventionist theory (presuppositions)

(a) The causal relation itself is not part of our ontology (**anti-realism**).

(b) **General causation** is the primary analysandum.

(c) **Manipulability relations** among event types are the primary analysans.

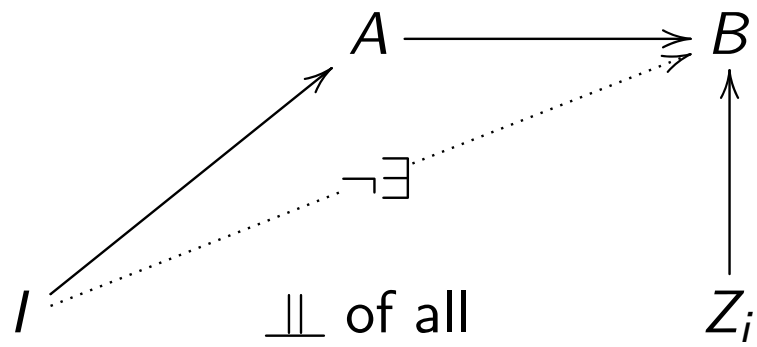
(d) Causation is deterministic.

(e) Causation is non-transitive.

Interventionist theory (conceptual preliminaries)

Intervention: I is an intervention on A with respect to B iff

1. I is causally relevant to A;
2. I is not connected to B on a path that does not go through A;
3. I is statistically independent of all causes Z_i of B that are not located on a path through A.



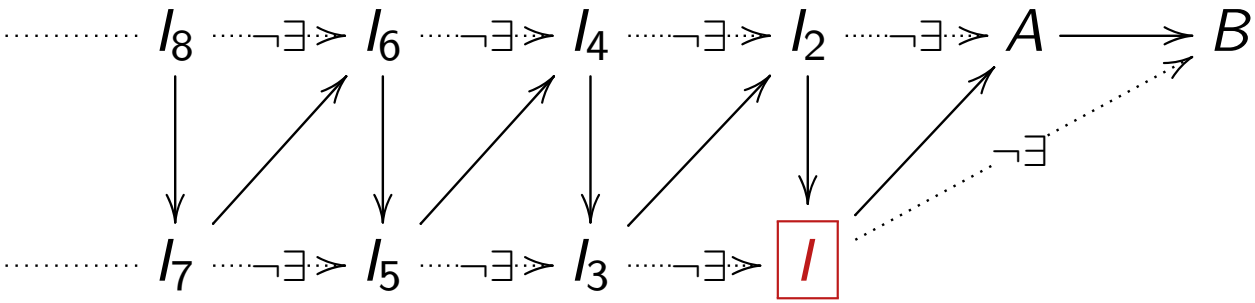
Interventionist theory (Woodward)

VI

A is causally relevant to *B* iff there exists a possible intervention *I* on *A* with respect to *B* that is accompanied by a change in *B* when all off-path causes of *B* are held fixed by interventions.

(VI) faces the following difficulties:

- (VI) is non-reductive and triggers an infinite regress:



- It is not clear in what sense interventions must be "possible".

Summary

- Causal notions are often used ambiguously and not always consistently.
- On the one hand, causes and effects shall be localized in time and space, on the other, absences and omissions shall also be causally interpretable.
- On the one hand, causal influence is believed to progress from one link of a causal chain to the next, on the other, a dislodged boulder is not taken to be a cause of a hiker's survival.
- On the one hand, causes shall determine their effects and events are believed not to occur without a cause, on the other, there shall be room for irreducibly indeterministic causal processes.

Summary

- Before a philosophical analysis of causation can get off the ground a consistent analysandum must be specified.
- This can only be done relative to a superordinate research goal or project. There is no correct and incorrect way to specify the analysandum.
- With the specification of the characteristics of the causal notion(s) to be analyzed the ensuing analysis is directed on a specific path.
- None of the known candidate theories of causation is free of problems and open questions.

Armstrong, D. M. (1983).
What is a Law of Nature?
Cambridge: Cambridge University Press.

Baumgartner, M. (2009).
Uncovering deterministic causal structures: A Boolean approach.
Synthese 170, 71–96.

Collins, J., N. Hall, and L. Paul (Eds.) (2004).
Causation and Counterfactuals, Cambridge. MIT Press.

Dowe, P. (2000).
Physical Causation.
Cambridge: Cambridge University Press.

Hall, N. (2000).
Causation and the price of transitivity.
Journal of Philosophy 97, 198–222.

Hitchcock, C. (2001).

The intransitivity of causation revealed in equations and graphs.
Journal of Philosophy 98, 273–299.

Hume, D. (1999 (1748)).
An Enquiry Concerning Human Understanding.
Oxford: Oxford University Press.

Kvart, I. (2004).
Probabilistic and counterfactual analyses.
In J. Collins, N. Hall, and L. A. Paul (Eds.), *Causation and Counterfactuals*, pp. 359–386. Cambridge: MIT Press.

Lewis, D. (1973).
Causation.
Journal of Philosophy 70, 556–567.

Lewis, D. (2000).
Causation as influence.
Journal of Philosophy 97, 182–197.

Mackie, J. L. (1974).

The Cement of the Universe. A Study of Causation.
Oxford: Clarendon Press.

Mill, J. S. (1843).
A System of Logic.
London: John W. Parker.

Pearl, J. (2000).
Causality. Models, Reasoning, and Inference.
Cambridge: Cambridge University Press.

Spirtes, P., C. Glymour, and R. Scheines (2000).
Causation, Prediction, and Search (2nd ed.).
Cambridge: MIT Press.

Suppes, P. (1970).
A Probabilistic Theory of Causality.
Amsterdam: North Holland.

Tooley, M. (1987).
Causation.

Oxford: Clarendon Press.

Woodward, J. (2003).
Making Things Happen. A Theory of Causal Explanation.
Oxford: Oxford University Press.