Analysing Causation

Dean McHugh

Institute of Logic, Language and Computation
University of Amsterdam

Causal Inference Lab
Master of Logic project
15 January 2020

Institute for Logic, Language and Computation

NWO

University of Amsterdam

Outline

1. Truth conditions for causal claims
   - Causation as probability raising
   - Causation as counterfactual dependence
   - Transitivity
   - Analysing causation via linguistic form

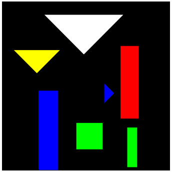
2. Graphical models of causation
   - Structural causal models, intuitively
   - Structural causal models, formally
   - Truth conditions of causal claims using SCMs

Help out with an experiment and win!

Sandro Pezzelle's experiment
Time slots available:

- Thursday 16 January 15:30–17:30
- Friday 17 January 11:00–13:00

*Is the red rectangle a **"big"** rectangle?*



We're collecting **answers by speakers** to this and similar questions — to explore the use and interpretation of **size adjectives**!

The experiment consists in playing a simple question-answering **game** (~20-40 mins) here at ILLC

Would you like to participate and win up to 7€?

If so, please send send an email to:
s.pezzelle@uva.nl

Plan

1. Truth conditions for causal claims
   - Causation as probability raising
   - Causation as counterfactual dependence
   - Transitivity
   - Analysing causation via linguistic form
2. Graphical models of causation
   - Structural causal models, intuitively
   - Structural causal models, formally
   - Truth conditions of causal claims using SCMs

Is causality all just probability?

- Given binary variables X and Y, say
  X raises the probability of Y just in case $P(y \mid x) > P(y)$.

Probability-raising is symmetric.

X raises the probability of Y iff Y raises the probability of X.

Proof.

$$P(y \mid x) > P(y)$$

$$\frac{P(x \mid y)P(y)}{P(x)} > P(y) \qquad \text{Bayes rule}$$

$$P(x \mid y) > P(x) \qquad \times \frac{P(x)}{P(y)}$$

□

∴ Probability-raising does not represent causal asymmetry

Example

Seeing someone smoke raising the probability that they cough

is equivalent to

Seeing someone cough raising the probability that they smoke

The asymmetry of time to the rescue

- Yes, probability-raising is symmetric
- But probability-raising that *follows the order of time* is not symmetric

Adding in time asymmetry

C causally contributed to E just in case

1. $P(E \mid C) > P(E)$, and
2. C happened before E

Note it does not matter whether we use $P(E \mid C) > P(E)$ or $P(E \mid C) > P(E \mid \neg C)$, as the two formulations turn out to be equivalent.

Probability as an objective ground for causation?

> *Probability is the most important concept in modern science, especially as nobody has the slightest notion what it means.*
>
> — Bertrand Russell, 1929 Lecture
> cited in Bell (1945, 587)

- The Kolmogorov axioms determine what constraints a probability distribution must satisfy (Kolmogorov, 1933)
- But they do not tell us what makes probabilistic claims true
- Two main approaches:

Subjective: Probability is one's degree of belief in a proposition
  - Associated with Bayesian statistics

Objective: Probability is the proportion of trials in the limit in which the outcome happens, if the trial were run infinitely often
  - Associated with frequentist statistics

Probability as an objective ground for causation?

- Many philosophers and statisticians favour a subjective interpretation of probability
- Some of these people would also want to say that (at least some) causal facts are mind-independent

Questions

1. Can one consistently hold:
    - A subjective interpretation of probability, and
    - An objective view of causation?
2. If so, what does this mean for probabilistic analyses of causation?

Plan

1. Truth conditions for causal claims
   - Causation as probability raising
   - Causation as counterfactual dependence
   - Transitivity
   - Analysing causation via linguistic form

2. Graphical models of causation
   - Structural causal models, intuitively
   - Structural causal models, formally
   - Truth conditions of causal claims using SCMs

Counterfactual analyses of causation: Lewis (1973)

> *We may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by object similar to the second. Or, in other words, where, if the first object had not been, the second never had existed.*
>
> — David Hume, *Enquiry Concerning Human Understanding*

> *[...] 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.*
>
> — David Lewis 'Causation' 1973

Causation as a chain of counterfactual dependence

Lewis (1973): C caused E just in case there is a chain of counterfactual dependence from C to E

Let's look at two problems for counterfactual analyses of causation:

1. Transitivity
2. Analysing causation via linguistic form

Plan

1. Truth conditions for causal claims
   - Causation as probability raising
   - Causation as counterfactual dependence
   - Transitivity
   - Analysing causation via linguistic form

2. Graphical models of causation
   - Structural causal models, intuitively
   - Structural causal models, formally
   - Truth conditions of causal claims using SCMs

Is causation transitive?

That causation is, necessarily, a transitive relation on events seems to many a bedrock datum, one of the few indisputable a priori insights we have into the workings of the concept.

— Ned Hall (2000)

- Lewis (1973) defines causation as the transitive closure of counterfactual dependence
- This makes causation transitive by definition
  - If C caused D and D caused E, then C caused E

Counterexamples to transitivity

Example (Boulder)

A boulder is dislodged, and begins rolling ominously toward Hiker. Before it reaches him, Hiker sees the boulder and ducks. The boulder sails harmlessly over his head. Hiker survives his ordeal.

Example (Dog bite)

Terrorist, who is right-handed, must push a detonator button at noon to set off a bomb. Shortly before noon, he is bitten by a dog on his right hand. Unable to use his right hand, he pushes the detonator with his left hand at noon. The bomb duly explodes.

See Halpern (2016, §2.4) and Beckers and Vennekens (2017) for discussion of transitivity

Plan

1. Truth conditions for causal claims
   - Causation as probability raising
   - Causation as counterfactual dependence
   - Transitivity
   - Analysing causation via linguistic form
2. Graphical models of causation
   - Structural causal models, intuitively
   - Structural causal models, formally
   - Truth conditions of causal claims using SCMs

Analysing causation via a linguistic form

- A counterfactual is a particular linguistic form
  - Exhibiting what Sabine Iatridou and Kai von Fintel have called *X*-marking (see, e.g., Iatridou and von Fintel, 2017)
- Counterfactual analyses of causation:
  reduce the truth conditions of causal claims to the truth conditions of another linguistic form (e.g. one exhibiting *X*-marking)

Analysing causation via another linguistic form

There is some sentence $S(x, y)$ with free variables x and y, such that, for any context c, "A caused B" is true in c just in case "$S(A, B)$" is true in c.

(A possible way out: specify what is meant by "counterfactual" without appealing to a particular linguistic form.)

Do counterfactuals have to track causal relations?

Example (Treasure hunt)

Ali and Bob are playing as a team in a treasure hunt, with one prize. A parent gives them a hint: *The prize is in the attic or the garden.*

Ali You go check the garden.
I'll search the attic.

Ali [*Later.*]
I found the prize in the attic!

Bob Well then why did you tell me to search in the garden?

Ali Because if the prize hadn't been in the attic, it would have been in the garden.

Example from Edgington (2011, p. 238), borrowed and adapted from Grice (1975)

Do counterfactuals have to track causal relations?

Is Ali's utterance true in the context of the treasure hunt?

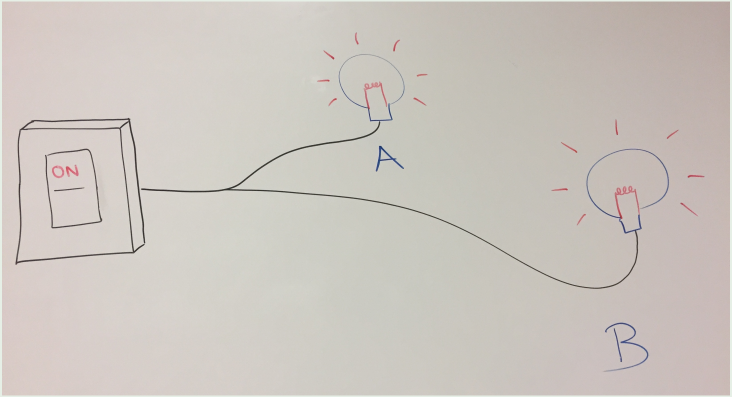
If the prize hadn't been in the attic,
it would have been in the garden.

- Edgington (2011, p. 239): "counterfactuals are too wide a class to hope to capture causation in terms of them"
- **Moral:** We should not analyse causation using the truth of a particular linguistic construction
- There is no guarantee that a particular linguistic form tracks causal relations (and not, e.g., epistemic relations)

Figure: Dorothy Edgington

Backtracking interpretations of counterfactuals

Example (Two lights)



1. If light *A* were off, light *B* would be off.
2. If light *A* turned off, light *B* would turn off.

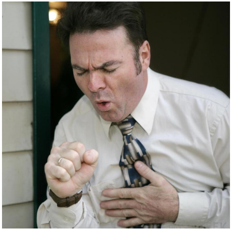
Plan

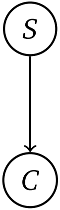
1. Truth conditions for causal claims
   - Causation as probability raising
   - Causation as counterfactual dependence
   - Transitivity
   - Analysing causation via linguistic form

2. Graphical models of causation
   - Structural causal models, intuitively
   - Structural causal models, formally
   - Truth conditions of causal claims using SCMs

Asymmetry by intervention

Observation

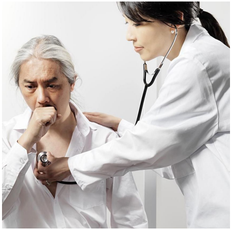




$P(c \mid s)$

Seeing someone cough
raises one's credence that they smoke

Intervention

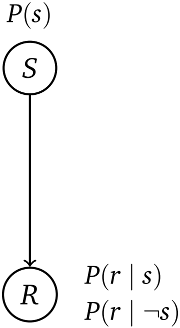


$P(c)$

Making someone cough
does *not* raise one's credence that they smoke

Structural causal model

- Edges in the graph represent direct causal dependence



The rooster's crow does not cause sunrise

Plan

Graphs

- Let W be a logical space (a set of possible worlds / states / situations / etc.)
- A variable X is a partition of W
- Where $x \in X$, we call x a value of X, and "$X = x$" an atomic sentence that is true at $w \in W$ just in case $w \in x$.
- Where X is a variable, $\mathcal{R}(X)$ is called the *range* of X, and denotes the set of values X may take
  - And where U is a set of variables, $\mathcal{R}(U)$ is defined as $\prod_{X \in U} \mathcal{R}(X)$

Definition (Graph)

A *graph* is a pair (V, E) where

- V is a set of variables
- $E \subseteq V \times V$ is a binary relation over V

Some graph terminology

- (V,E) is *acyclic* just in case E is acyclic
  - Where E^+ is the transitive closure of E, E is *acyclic* iff for no $X \in V$ is $(X,X) \in E^+$
- There is a *directed path* from X to Y in (V,E) iff $(X,Y) \in E^+$
- A graph is a family!
  - X is a *parent* of Y, and Y is a *child* of X, iff $(X,Y) \in E$
    - $pa_X := \{Y \in V : (Y,X) \in E\}$
  - X is an *ancestor* of Y, and Y is a *descendent* of X, iff $(X,Y) \in E^+$
- A variable is *exogenous* iff it has no parents, and *endogenous* otherwise

Structural causal models

Definition (Structural causal model)

A structural causal model is a triple $M = (V, E, F)$ where

- V is a set of variables
- (V, E) is a directed acyclic graph
- F is a set of functions of the form

$$F_X : \mathcal{R}(pa_X) \to \mathcal{R}(X),$$

  one for each endogenous variable $X \in V$

The value of an endogenous variable X is determined by the values of its parents, according to F_X

- Since F_X are functions, the dependence is deterministic
- Where $U = u$ is an assignment of values to the exogenous variables in V, we call u a *setting* or *context* for M
  - I.e. the values of the exogenous variables determine the values of all the variables

Interventions in structural causal models

Let $M = (V, E, F)$ be a structural causal model

Definition (Interventions as model surgery)

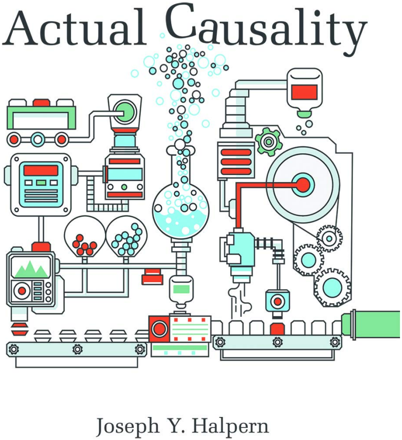
$M_{X=x}$ is the model $(V, E, F_{X=x})$ which results from replacing the equation for X in M with $X = x$ (that is, $F_{X=x} := (F \setminus \{F_X\}) \cup \{F'_X\}$ where $F'_X(y_1, y_2, \ldots) = x$ for any values $y_1, y_2, \ldots$ of X's parents).

Definition (Truth conditions for interventions)

Let M be a structural causal model and u a setting of the exogenous variables.

$$M, u \models [X \leftarrow x]Y = y \quad \text{iff} \quad M_{X=x}, u \models Y = y$$

Joseph Halpern, *Actual Causality* (2016)



Plan

1. Truth conditions for causal claims
   - Causation as probability raising
   - Causation as counterfactual dependence
   - Transitivity
   - Analysing causation via linguistic form
2. Graphical models of causation
   - Structural causal models, intuitively
   - Structural causal models, formally
   - Truth conditions of causal claims using SCMs

Halpern (2016)

Definition (Truth conditions for actual causal claims)

$X = x$ is an *actual cause* of φ in the causal setting (M, u) iff

- AC1 $(M, u) \models X = x$ and $(M, u) \models \varphi$
- AC2 There is a set W of variables and a setting x' of the variables in X such that, if $(M, u) \models W = w$ then $(M, u) \models [X \leftarrow x', W \leftarrow w]\neg\varphi$
- AC3 X is minimal; there is no strict subset X' of X such that $X' = x'$ satisfies conditions AC1 and AC2, where x' is the restriction of x to the variables in X'

- AC1 The cause and effect actually occurred
- AC2 Fixing some variables to their actual values, there is a value of the cause x' such that the effect would not have occurred

Keeping things fixed

A technical modification of AC2:

AC2′ There is a set W of variables and a setting x' of the variables in X such that $(M,u) \models W = w$ and

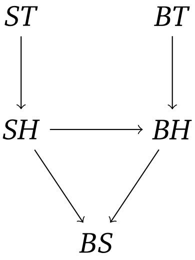
$$(M,u) \models [X \leftarrow x', W \leftarrow w]\neg\varphi$$

- Braham and van Hees (2012)'s definition of causation: one can fix actual facts
- This allows one to assign causal contributions under both disjunctive and conjunctive voting rules
- Halpern (2016)'s definition does the same

Overdetermination

Example (Late preemption)

Suzy and Billy both throw a rock at a bottle. Suzy's rock gets there first, shattering the bottle. However Billy's throw was also accurate, and would have shattered the bottle had it not been preempted by Suzy's throw.



$$SH = ST$$
$$BH = BT \wedge \neg SH$$
$$BS = SH \vee BH$$

Figure: Halpern's model of *Late preemption* (2016, p. 31)

Beckers' discussion of Halpern (2016)

Sander Beckers (2016)

- Halpern makes the right prediction, but for the wrong reason
  - **Right prediction:** Suzy, and not Billy, caused the bottle to break
  - **Wrong reason:** This is not because we imagine, if Suzy had not thrown, but Billy's rock had still not hit the window, that the window would not have broken
- "[The scenario] cannot be interpreted simply as the possibility that the backup mechanism fails to function properly, because the actual story explicitly stipulates that it does not" (Beckers and Vennekens, 2018, p. 853)
- Instead, we have to take the timing into account:
  - Billy's throw did not cause the bottle to break because the bottle actually broke before Billy's rock would have hit it

References I

Sander Beckers. *Actual Causation: Definitions and Principles*. PhD thesis, PhD thesis, KU Leuven, 2016. URL `https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1656621&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US`.

Sander Beckers and Joost Vennekens. The transitivity and asymmetry of actual causation. *Ergo: An Open Access Journal of Philosophy*, 4:1–27, 2017. doi:10.3998/ergo.12405314.0004.001.

Sander Beckers and Joost Vennekens. A principled approach to defining actual causation. *Synthese*, 195(2):835–862, Feb 2018. doi:10.1007/s11229-016-1247-1.

Eric Temple Bell. *The Development of Mathematics, 2nd edition*. McGraw-Hill Book Company, 1945.

Matthew Braham and Martin van Hees. An anatomy of moral responsibility. *Mind*, 121 (483):601–634, 7 2012. doi:10.1093/mind/fzs081.

Dorothy Edgington. Causation first: Why causation is prior to counterfactuals. In Christoph Hoerl, Teresa McCormack, and Sara R. Beck, editors, *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology*, chapter 11, pages 230–241. Oxford University Press, 2011.

References II

Herbert P Grice. Logic and conversation. In *Speech acts*, pages 41–58. Brill, 1975.

Ned Hall. Causation and the price of transitivity. *Journal of Philosophy*, 97(4):198, 2000. doi:10.2307/2678390.

Joseph Y Halpern. *Actual Causality*. MIT Press, 2016.

Sabine Iatridou and Kai von Fintel. Transparent wishes. Lecture slides, 2017. URL `http://web.mit.edu/fintel/fintel-iatridou-2017-x-desires-slides-ipac.pdf`.

A. N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitrechnung,* Ergebnisse Der Mathematik; translated as *Foundations of Probability*. Chelsea Publishing Company, 1950, 1933.

David Lewis. Causation. *Journal of Philosophy*, 70(17):556–567, 1973. doi:10.2307/2025310.