Actual causality in dynamical systems

Dean McHugh

Institute of Logic, Language and Computation
University of Amsterdam

Statistical inference workshop, ILLC, UvA
3 December 2019

Goal: analyse the logical structure of **hypothetical reasoning**

- Planning
- Decision-making
- Moral responsibility; praise, blame
- Causality

Cases when our reasoning goes beyond actuality

Case study: glyphosate

Figure: Protests outside European Commission, 24 October 2017

The New York Times

Report Finds Traces of a Controversial Herbicide in Cheerios and Quaker Oats

Glyphosate is a popular weedkiller used on crops worldwide, but it has been at the center of a debate over its presence in foods. A new report found traces in Cheerios, Quaker Oats and other breakfast foods. Seth Perlman/Associated Press

By **Mihir Zaveri**

Aug. 15, 2018

Monsanto banned from European parliament

MEPs withdraw parliamentary access after the firm shunned a hearing into allegations that it unduly influenced studies into the safety of glyphosate used in its RoundUp weedkiller

Arthur Neslen

Thu 28 Sep 2017 15.15 BST

5,833

▲ People protest against a planned $66bn takeover of Monsanto by Bayer and Monsanto's glyphosate herbicides, in Brussels, July 2017. Photograph: Yves Herman/Reuters

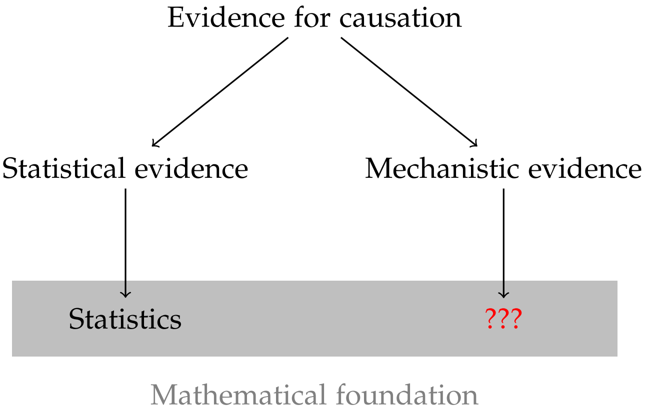
ECHA
EUROPEAN CHEMICALS AGENCY

European Chemicals Agency

- Only statistical evidence is valid
  - Epidemiological data
  - Studies on animals
  - Mostly conducted by industry itself

World Health Organisation

- "The mechanistic data provide strong evidence" that glyphosate causes cancer
- But what is *mechanistic evidence*?


Evidence for causation
Statistical evidence
Mechanistic evidence
Statistics
???
Mathematical foundation

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.

$$
\begin{aligned}
P(y \mid x) &> P(y) \\
\frac{P(x \mid y)P(y)}{P(x)} &> P(y) && \text{Bayes rule} \\
P(x \mid y) &> P(x) && \times \tfrac{P(x)}{P(y)}
\end{aligned}
$$

□

∴ Probability-raising does not represent causal asymmetry

Example

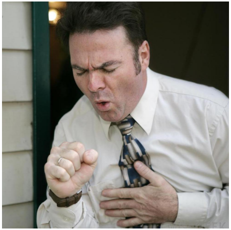
Seeing someone smoke raising the probability that they cough

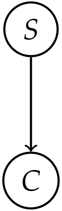
is equivalent to

Seeing someone cough raising the probability that they smoke

Asymmetry by intervention

Observation

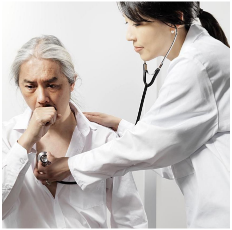




$P(c \mid s)$

Seeing someone cough raises one's credence that they smoke

Intervention

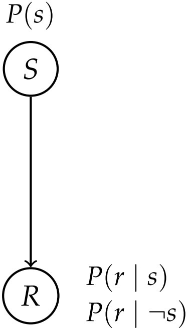


$P(c)$

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

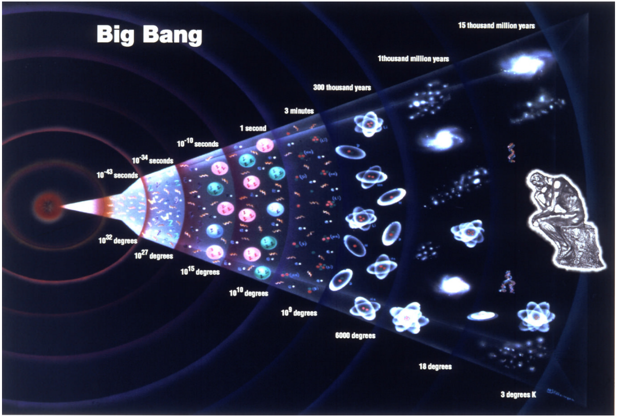
Bayes net

- Edges in the graph represent direct causal dependence.



The rooster's crow does not cause sunrise

Is causality all just intervention?



- There are causal relations at the big bang
- But what would it mean to *intervene* in the big bang?
- Intervention is too anthropocentric

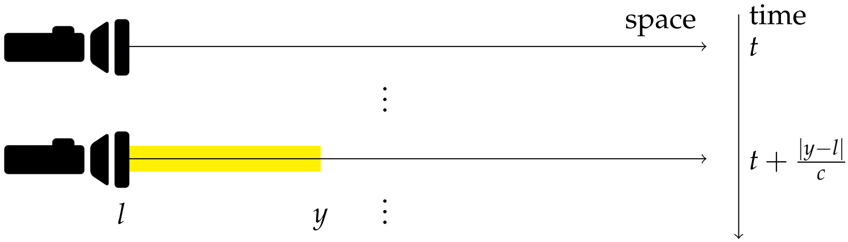


Figure: Turning on the light at time t

Fact

No binary relation is nonempty, dense, and anti-transitive.

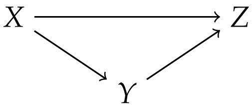


Figure: Proof

Definition (Dense causal chain)

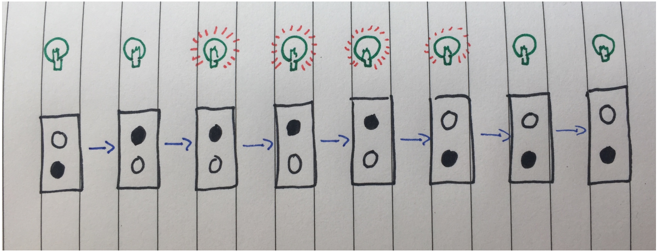
A Bayesian net has a *dense causal chain* just in case it contains variables X and Z such that:

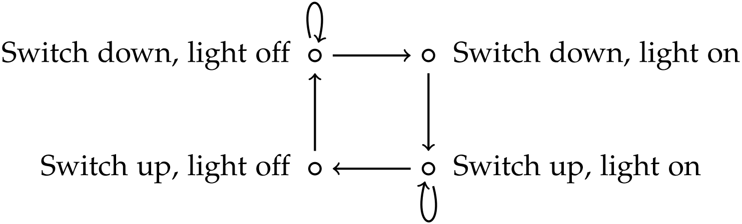
1. X has a causal effect on Z $\quad P(z \mid do(x)) \neq P(z)$
2. X is a parent of Z
3. There is another parent Y of Z such that X and Z are independent conditional on any set containing Y

Theorem

No Bayesian net has a dense causal chain.

Temporal modality





Dynamical systems

Let S be a set of states (e.g. 'worlds', 'situations', 'valuations').

Definition (Dynamical system)

- A *path* is a linearly ordered set of states (where we allow the same state to appear multiple times).
- A *dynamical system* is a set of paths.

- A dynamical system encodes what paths are *possible*
- Flexible definition: e.g. the linear order can be dense

Illustration: and-gate

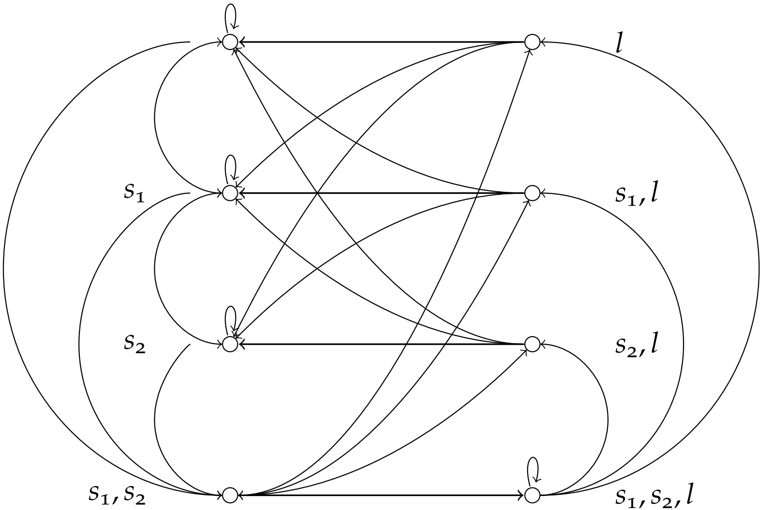


Figure: Dynamical system representing an and-gate

Probability in dynamical systems

Given a joint probability distribution P, how to calculate P_s?

- **Strategy 1** Uniform across states

$$P_s(t) := P(t)$$

- **Strategy 2** Inspired by the chain rule
  - Order the variables $X_1, \ldots, X_n$
  - Let $s(X_i)$ be the value of X_i at s

Static chain rule

$$P(s) = \prod_{i=1}^{n} P(s(X_i) \mid s(X_1), \ldots, s(X_{i-1}))$$

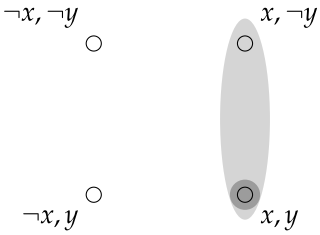
Dynamic chain rule

$$P_s(t) := \prod_{i=1}^{n} P(t(X_i) \mid s(X_1), \ldots, s(X_{i-1}))$$

Two kinds of conditionalization

Static

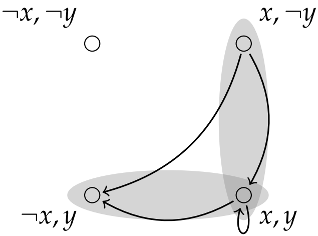
- About simultaneous information
- $P(y \text{ now} \mid x \text{ now})$



$$P(y \mid x) = \frac{P(x,y)}{P(x)} = \sum_{s \Vdash x,y} \frac{P(s)}{P(x)}$$

Dynamic

- About changing information
- $P(y \text{ next} : x \text{ now})$



$$P(y : x) = \sum_{s \Vdash x} P_s(y) \frac{P(s)}{P(x)}$$

Does this even work?

Theorem

Let $\mathcal{M}$ be a discrete dynamical system, with state $s \in S$, and $P : S \to [0, 1]$ a probability distribution. For any $t \in S$ put $P_s(t) = P(t : s)$. Then P_s is a probability distribution.

Proof

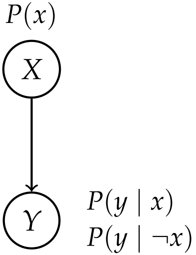
Theorem

Let $\mathcal{M}$ be a discrete dynamical system, with state $s \in S$, and $P : S \to [0, 1]$ a probability distribution. For any $t \in S$ put $P_s(t) = P(t : s)$. Then P_s is a probability distribution.

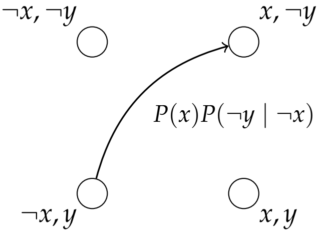
Proof.

Since P is a probability distribution, $P_s : S \to \mathbb{R}$ and $P_s(t) \geq 0$ for any $t \in S$. We show that $1 = \sum_{t \in S} P(t : s)$ by induction on the number of variables $X_1, \ldots, X_n$. Let $V(X) = \{v_1(X), \ldots v_k(X)\}$ be the values a variable X can take. For $n = 1$, $\sum_{t \in S} P(t : s) = \sum_{t \in S} P(t(X_1) \mid s(pa(X_1)))$. As atom-equivalent states are identical, $\sum_{t \in S} P(t(X_1) \mid s(pa(X_1))) = \sum_{v(x_1) \in V(x_1)} P(v(X_1) \mid s(X_1))$, which is 1 since P is additive and the values of X_1 are mutually exclusive. For $n = m + 1$, $\sum_{t \in S} P(t : s) = \sum_{t \in S} \prod_i^n P(t(X_i) \mid s(pa(X_i))) = \sum_{t \in S} \left(P(t(X_n) \mid s(pa(X_n))) \prod_n^m P(t(X_i) \mid s(pa(X_i)))\right) = \sum_{v(X_n) \in V(X_n)} \left(P(v(X_n) \mid s(pa(X_n))) \sum_{t \in S} \prod_i^m P(t(X_i) \mid s(pa(X_i)))\right)$, which by induction hypothesis is $\sum_{v(X_n) \in V(X_n)} P(v(X_n) \mid s(pa(X_n)))$ which is 1 again by additivity of P and exclusivity of each $v(X_n)$. Hence $P_s(t) \leq 1$, so $P_s : S \to [0, 1]$. □

Bayes nets as dynamical systems



Bayes Net [see 1, 2]



As a dynamical system

- Graph gives the order $P(s) = \prod_{s \Vdash v} P(v \mid pa(v))$
- $P_s(t) = \prod_{\substack{s \Vdash pa(v) \\ t \Vdash v}} P(v \mid pa(v))$
- Update $[!x]$: delete $\neg x$-states, condition on x
- Intervene $do(x)$: make every state an x-state

Intervention in dynamical systems

If $s \Vdash a$ then

$$P(s \mid observe\ a) = \frac{P(s)}{P(a)}$$

$$P(s \mid do\ a) = \frac{P(s)}{P_s(a)}$$

If $s \nVdash a$ then $P(s \mid observe\ a) = P(s \mid do\ a) = 0$.

- Delete the $\neg a$-states, then
    - Observation: $P_s(t)$ becomes $P_s(t \mid a)$
    - Intervention: $P(s)$ becomes $P_s(t) + P_s(t^a)$

where t^a is t's a-counterpart, agreeing with t on the value of every variable but a.

From Bayesian nets to dynamical systems

Let B, B' be two Bayesian nets with the same set of variables V.

- Let $\mathcal{D}(B)$ be the representation of B a dynamical system.
- Let us call B and B' *interventionally equivalent* ($B \sim_{do} B'$) just in case

$$P_B(\vec{X} \mid do(\vec{Y})) = P_{B'}(\vec{X} \mid do(\vec{Y}))$$

for all sets of variables $\vec{X}, \vec{Y}$ in V.

Theorem (Representation theorem)

If $B \sim_{do} B'$ then $\mathcal{D}(B) = \mathcal{D}(B')$.

Inevitable effects

- Intuitive observation about causality: effects depend on their causes
- The effect occurring rather than not occurring depends on the cause occurring rather than not occurring
- ⇒ The effect might not have occurred

What about inevitable events?

(1) Socrates drinking poison caused him to die.

What we talk about when we talk about events

(2) Drinking poison actually caused Socrates to die, but that was always going to happen at some point.

(3) Socrates was going to die eventually. Plato's *Phaedo* says that it was actually caused by drinking poison.

(4) Socrates drinking poison caused the occurrence of an event that was bound to occur ever since he was born; namely, his death.

Token dependence without causation

Switch. *An engineer is standing by a switch in the railroad tracks. A train approaches in the distance. She flips the switch, so that the train travels down the left-hand track, instead of the right. Since the tracks reconverge up ahead, the train arrives at its destination all the same.*

(5) The engineer flipping the switch did not cause the train to reach its destination

Causation without token dependence

Parliament. *A parliament voted on a bill which only required a single vote to pass (the bill is a mere formality, say). Passed bills are signed into law by the President every year on January 1st. Suppose that two members of parliament, Alice and Bob, voted for the bill, so it passed and on January 1st become law.*

(6) Alice's vote for the bill caused it to become law, and Bob's vote for the bill caused it to become law.

Beckers (2016)

Beckers (2016) definition of production

1. Each event on the chain actually occurred
2. Each event on the chain occurred before the next in time
3. For each C_i on the chain we can find a set of facts L in the actual scenario such that each fact in L occurred before C_{i+1}, and C_i on its own is not sufficient for C_{i+1} but $L \cup \{C_i\}$ is sufficient for C_{i+1}

- For C to produce E, there must be some set of facts that are not sufficient for E
  - There must be a context (i.e. assignment of values to variables) in which E does not occur

Upshot: Inevitable events have no producers.

Is causation without any kind of dependence possible?
What would it look like?

What is causality about?

The ways the world can change through time

- Causality is a concept of temporal modality,
  - a logical pattern in time

Dynamical systems

- Ingredients
  - A set of worlds W
  - A set of times $\mathcal{T}$ with a linear order $\leq$
- Worlds are temporally extended
- Objects of evaluation
  - World-time pairs $\langle w, t \rangle$ ($w \in W$ and $t \in \mathcal{T}$)
  - $\langle w, t \rangle$ the world w at time t
  - We will call $\langle w, t \rangle$ a state of w, and often write it as w_t
- Primitive notion of the theory: truth at a world-time pair
  - Truth for instantaneous information

Preliminary observations about causation

- Causes 'generate', 'bring about', 'produce' their effects
  - i.e. with respect to some background assumptions, causes are sufficient for their effects
- Causation involves a continuous chain of events from the cause to the effect

Sufficiency

- Sufficiency = "Whenever C occurs, E will occur"

Definition (Global sufficiency)

C is *sufficient* for E just in case for any world w and time t such that C is true at w_t, there is a later time $t' > t$ such that E is true at $w_{t'}$.

- Background assumptions are restrictions on worlds

Definition (Sufficiency)

Let A be a set of worlds. C is *sufficient* for E given A just in case for any world w in A and time t such that C is true at w_t, there is a later time $t' > t$ such that E is true at $w_{t'}$.

Chains of events

We have an intuitive picture that causation involves a chain of events from the cause to the effect.

- Where C is a sentence, t a time and $w_{t'}$ a state, C_t is true at $w_{t'}$ just in case C is true at $w_{t'}$ and $t = t'$
  - C is the event-type, and C_t a token event
  - C_t is an particular instance of C's occurrence
- Causal chains are chains of token events

Definition (Chain of events)

Where T is a set of times, and we call $\{C_t\}_{t \in T}$ a *chain of events*.

- (Notation: we allow the C's in $\{C_t\}_{t \in T}$ to be distinct.)

Continuity

Late preemption. *Billy and Suzy throw a rock at a bottle. Both throws are accurate, but Suzy's rock hits the bottle first, breaking it. If Suzy's rock hadn't hit the bottle, Billy's rock would have, and the bottle would still have broken.*

- Suppose we are using chains of events to analyse causation
- but we allow chains of events between causes and effects to have gaps
  - i.e. there is a time t between the time when the cause occurred and the time when the effect occurred, with no event on the chain occurring at t
- If we leave out the time including when Suzy hit the bottle, the case above would be indistinguishable from one where Billy caused the bottle to break

Chains of events with gaps do not track causal relations.

Density

Definition (Density)

A set of times T is *dense* just in case for any $t, t' \in T$ and time t^*

$$\text{if } t < t^* < t' \text{ then } t^* \in T.$$

A chain of events $\{C_t\}_{t \in T}$ is *dense* just in case T is dense.

Question: Can we plug the gaps with any events whatsoever?

- There has to be some relation between events on the chain
  - Besides occurring one after another in time
- Each event on the chain should be sufficient for the next
  - When time is dense, the concept of a unique "next" time is not defined

Sufficiency

Sufficiency is transitive

- If C is sufficient for D and D is sufficient for E, then C is sufficient for E
- Each event on a causal chain is sufficient for the next ✗
- Each event on a causal chain is sufficient for the later events ✓
  - Since sufficiency is transitive, the latter does not require more than the former, but is defined for dense paths

Definition (Sufficiency-preserving)

Where A is a set of worlds, a chain of events $\{C_t\}_{t \in T}$ is *sufficiency-preserving* given A just in case for any $t, t' \in T$,

$$\text{if } t < t' \text{ then } C_t \text{ is sufficient for } C_{t'} \text{ given } A.$$

From the cause to the effect

- Given sentences C and E, a chain of events is a chain *from C to E* just in case C's occurrence is the first event on the chain and E's occurrence is the last
  - That is, when $\{D_t\}_{t \in T}$ is the chain of events, $C_{\min(T)} = D_{\min(T)}$ and $E_{\max(T)} = D_{\max(T)}$, where
    - $\min(T) = t$ iff $t \in T$ and $t \leq t'$ for all $t' \in T$
    - $\max(T) = t$ iff $t \in T$ and $t \geq t'$ for all $t' \in T$
- A chain of events exists at a world just in case each event on the chain occurs in that world

Production

Definition (Production)

C produced E at w_t just in case C is sufficient at w_t for the existence of some dense, sufficiency-preserving chain of events from C to E.

- C's occurrence (at some time) is sufficient for the existence of some chain token events such that
  - The first event on the chain is an instance of C
  - The final event on the chain is an instance of E
  - For each time between C's occurrence on the chain and E's occurrence on the chain, there is an event on the chain occurring at that time
  - The occurrence of each event on the chain, occurring at the particular time it occurs, is sufficient for every later event on the chain occurring at the particular time it occurs

Types & tokens

Production is a relation between event types, requiring the existence of a chain of event tokens.

Many ways to cause an effect

Definition (Production)

C produced E at w_t just in case C is sufficient at w_t for the realisation of some dense, sufficiency-preserving chain of events from C to E.

- This does not say that for C to produce E, there is a chain of events such that C is sufficient for its realisation

Many ways to cause an effect

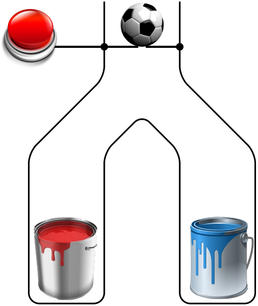


Figure: Two effects, one way to produce each

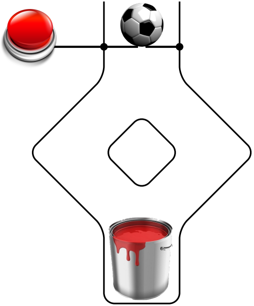


Figure: Two ways to produce one effect

From production to causation

Following Beckers (2016),

- With the addition of background assumptions

Definition (Actual causation)

Let A be a set of background assumptions.
C actually caused E with respect to A just in case

- C produced E given A
- If C had not occurred, but A had still occurred, $\neg C$ would have produced E given A

Preemption

> ***Late preemption.*** *Billy and Suzy throw a rock at a bottle. Both throws are on target, but Suzy's rock hits the bottle first, breaking it. If Suzy's rock had not hit the bottle, Billy's rock would have, and the bottle would still have broken.*

Suzy's throw caused the bottle to break

- Her throw was sufficient for a chain of events of the form: Suzy's rock flies through the air on target at time t

For what was Billy throwing the rock sufficient?

- Billy's rock flying on target through the air at t

Each of those events was not sufficient for the bottle to break when it did, but only at a later time

Inevitable effects revisited

- Socrates drinking poison caused his death because
  - Socrates drinking poison produced his death
  - If he hadn’t drank poison, the fact that he did not drink poison would not have produced his death