Difference-making

Philosophical foundations of explanation

Dean McHugh

Institute of Logic, Language and Computation
University of Amsterdam

January 10, 2024

Institute for Logic, Language and Computation

NWO

University of Amsterdam

Plan

- Sufficiency for production

1. Difference making
   - Sartorio's analysis of difference-making
   - Unravelling Sartorio's Principle
   - The ubiquity of the Perfection Principle
   - On the pragmatic origins of the Perfection Principle

Plan

- Sufficiency for production

1. Difference making
   - Sartorio's analysis of difference-making
   - Unravelling Sartorio's Principle
   - The ubiquity of the Perfection Principle
   - On the pragmatic origins of the Perfection Principle

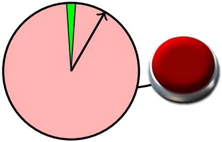
Sufficiency **and** production.

C *cause* E entails $(C \gg E) \wedge (C$ *produce* $E)$.

Sufficiency **for** production.

C *cause* E entails $C \gg (C$ *produce* $E)$.

Alice and Bob are two children at the funfair with their parents. The parents decide that the children should have a souvenir of their time there: if any child does not win a teddy by the end of the day, the parents will buy one for them.



Alice and Bob play a game with a spinner and a button. A pointer moves around the circle until the player pushes the button. If the pointer lands in the thin green region, the player wins a teddy. If it lands in the red region, the player gets nothing.

Alice entered the game and, by sheer luck, pushed the spinner at the right time. The pointer landed in the winning region and she won a teddy. Bob entered the game and pushed the button at the wrong time. The pointer landed in the red region and he didn't win a teddy.
At the end of the day, the parents notice that Bob didn't win a teddy, so they bought him one.

(1) Alice got a teddy because she entered the spinner game.

Alice entered the game and, by sheer luck, pushed the spinner at the right time. The pointer landed in the winning region and she won a teddy. Bob entered the game and pushed the button at the wrong time. The pointer landed in the red region and he didn't win a teddy.
At the end of the day, the parents notice that Bob didn't win a teddy, so they bought him one.

(1) Alice got a teddy because she entered the spinner game.

(2) Alice got a teddy because she won the spinner game (and Bob got a teddy because his parents bought him one).

Sufficiency: Alice enter $\gg$ get teddy ✓

Since Alice was guaranteed to get a teddy, anything whatsoever was sufficient for her to get a teddy. In particular, Alice entering the game was sufficient for her to get a teddy.

Production: Alice enter *produce* get teddy ✓

The fact that Alice entered the game produced her to get a teddy. Without providing a formalization of the scenario, we can loosely say that Alice entering the game produced her to get a teddy because there is a chain from her entering the game to her getting a teddy from the game such that each condition on the chain counterfactually depends on a buffer of previous conditions on the chain.

Sufficiency for production: Alice enter $\gg$ (Alice enter *produce* get teddy)

✗

Given the set up of the scenario, it was determined that Alice would get a teddy somehow, but it was not determined how she would get it. She could have hit the button at the wrong time and lost the spinner game (a possibility emphasized by the presence of Bob), in which case she would have gotten a teddy from her parents at the end of the day rather than from winning the game. In that case her entering the game would not have produced her to get a teddy; rather, her parents buying a teddy would have produced her to get one. In all, then, the fact that Alice entered the game was not sufficient for her entering the game to produce her to win a teddy.

"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."

(Lewis 1973)

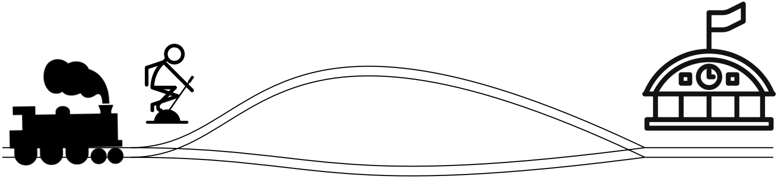


Figure: Switching scenario from Hall (2000, p. 205).

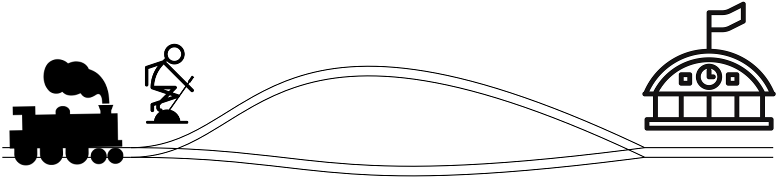


Figure: Switching scenario from Hall (2000, p. 205).

(3) a. The train reached the station because the engineer flipped the switch. ✗

b. The engineer flipping the switch caused the train to reach the station. ✗

C caused E
E because C

C caused E *E because C*	$\Rightarrow$	C made a difference to E

C *caused* E
E *because* C

$\Rightarrow$

C made a difference to E

Today's question

What does "C made a difference to E" mean?

Hypothesis: Difference making is counterfactual dependence

C made a difference to E

just in case

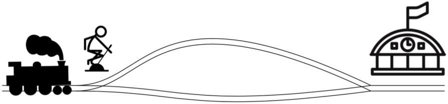
if C had not occurred, E would not have occurred.

Hypothesis: Difference making is counterfactual dependence

C made a difference to E

just in case

if C had not occurred, E would not have occurred.



One thing that catches the eye ... is that, just as the *flip* doesn't make a difference to the [train reaching the station], the *failure to flip* wouldn't have made a difference to the [train reaching the station] either. In other words, *whether or not* I flip the switch makes no difference [to the train's arrival], it only helps to determine the route that the train takes [to the station].

(Sartorio 2005, pp. 74–75)

Plan

- Sufficiency for production

1. Difference making
   - Sartorio's analysis of difference-making
   - Unravelling Sartorio's Principle
   - The ubiquity of the Perfection Principle
   - On the pragmatic origins of the Perfection Principle

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \ \textit{cause} \ E \qquad \Rightarrow \qquad \neg C > \neg(\neg C \ \textit{cause} \ E)$$

Sartorio's Principle accounts for the switch case:

- Suppose for reductio that the engineer pulling the lever caused the train to reach the station.

Sartorio's Principle accounts for the switch case:

- Suppose for reductio that the engineer pulling the lever caused the train to reach the station.
- Then (intuitively) if the engineer hadn't pulled the lever, that would have also caused the train to reach the station.

Sartorio's Principle accounts for the switch case:

- Suppose for reductio that the engineer pulling the lever caused the train to reach the station.
- Then (intuitively) if the engineer hadn't pulled the lever, that would have also caused the train to reach the station.
- Sartorio's principle is violated.

Sartorio's Principle accounts for the switch case:

- Suppose for reductio that the engineer pulling the lever caused the train to reach the station.
- Then (intuitively) if the engineer hadn't pulled the lever, that would have also caused the train to reach the station.
- Sartorio's principle is violated.
- Then assuming Sartorio's principle, the engineer pulling the lever did not caused the train to reach the station.

Sartorio's Principle also accounts for the Billy and Suzy case:

- Imagine that Suzy had not thrown.

Sartorio's Principle also accounts for the Billy and Suzy case:

- Imagine that Suzy had not thrown.
- In that case Billy's rock would have hit the window, and it would have broken anyway.

Sartorio's Principle also accounts for the Billy and Suzy case:

- Imagine that Suzy had not thrown.
- In that case Billy's rock would have hit the window, and it would have broken anyway.
- Intuitively, Billy's throw caused the window to break.

Sartorio's Principle also accounts for the Billy and Suzy case:

- Imagine that Suzy had not thrown.
- In that case Billy's rock would have hit the window, and it would have broken anyway.
- Intuitively, Billy's throw caused the window to break.
- But what about Suzy *not* throwing?
  Did that cause the window to break?
  Intuitively not!

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \ cause \ E \quad \Rightarrow \quad \neg C > \neg(\neg C \ cause \ E)$$

Sartorio's Principle gives us a principled way to distinguish Suzy and the switch.

- Suzy's throw satisfies Sartorio's Principle.
- Pulling the switch does not.

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \ cause \ E \quad \Rightarrow \quad \neg C > \neg(\neg C \ cause \ E)$$

Sartorio's principle is automatically satisfied when the effect counterfactually depends on the cause.

- Suppose $\neg C > \neg E$

Counterfactual dependence is one way to make a difference.

But, as Suzy shows, it is not the only way to make a difference.

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \textit{ cause } E \quad \Rightarrow \quad \neg C > \neg(\neg C \textit{ cause } E)$$

Sartorio's principle is automatically satisfied when the effect counterfactually depends on the cause.

- Suppose $\neg C > \neg E$
- $\neg E$ entails $\neg(\neg C \textit{ cause } E)$

Counterfactual dependence is one way to make a difference.

But, as Suzy shows, it is not the only way to make a difference.

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \textit{ cause } E \quad \Rightarrow \quad \neg C > \neg(\neg C \textit{ cause } E)$$

Sartorio's principle is automatically satisfied when the effect counterfactually depends on the cause.

- Suppose $\neg C > \neg E$
- $\neg E$ entails $\neg(\neg C \textit{ cause } E)$
- Hence $\neg C > \neg(\neg C \textit{ cause } E)$

Counterfactual dependence is one way to make a difference.

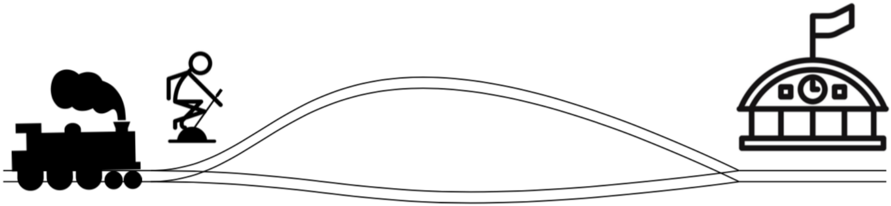
But, as Suzy shows, it is not the only way to make a difference.

Plan

- Sufficiency for production

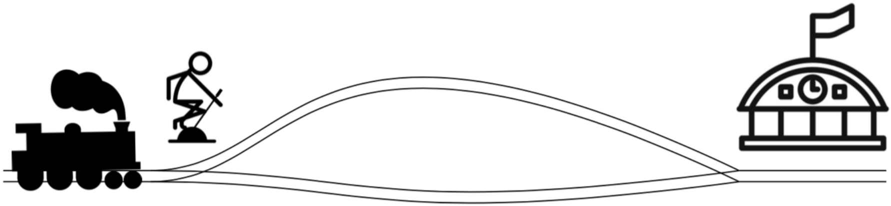
1. Difference making
   - Sartorio's analysis of difference-making
   - Unravelling Sartorio's Principle
   - The ubiquity of the Perfection Principle
   - On the pragmatic origins of the Perfection Principle

Suppose we have a semantics of *cause*, call it *proto-cause*, that does not account for the switches case.



Pulling the lever proto-caused the train to reach the station.

Suppose we have a semantics of *cause*, call it *proto-cause*, that does not account for the switches case.



Pulling the lever proto-caused the train to reach the station.

Challenge

How do we amend *proto-cause* to predict that pulling the lever did **not** cause the train to reach the station?

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \ cause \ E \qquad \Rightarrow \qquad \neg C > \neg(\neg C \ cause \ E)$$

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \ cause \ E \qquad \Rightarrow \qquad \neg C > \neg(\neg C \ cause \ E)$$

$$\begin{aligned} x &\geq 3x - 2 \\ x + 2 &\geq 3x \\ \frac{x+2}{3} - x &\geq 0 \end{aligned}$$

Sartorio's Principle

If C caused E, then, had C not occurred, the absence of C wouldn't have caused E.

$$C \ cause \ E \quad \Rightarrow \quad \neg C > \neg(\neg C \ cause \ E)$$

$$x \geq 3x - 2$$
$$x + 2 \geq 3x$$
$$\frac{x+2}{3} - x \geq 0$$

Problem

- The operations of arithmetic have inverses (addition/subtraction; multiplication/division)
- Logical operations do not

Definition

Let $A[C/B]$ be the result of replacing every occurrence of B in A with C.

Example

$$((p \vee q) \wedge \neg q)[r/q] \quad = \quad (p \vee r) \wedge \neg r.$$

The Perfection Principle.

For any sentences C and E, there is a sentence X such that C *cause* E entails $C > X$ and $\neg(C > X)[\neg C/C]$.

Sartorio's Principle

$C \ \mathit{cause} \ E \ \Rightarrow \ \neg C > \neg(\neg C \ \mathit{cause} \ E)$

The Perfection Principle

$C \ \mathit{cause} \ E \ \Rightarrow \ (C > X) \ \wedge \ \neg(C > X)[\neg C/C]$

Let $A \diamondto C$ abbreviate $\neg(A > \neg C)$.

(4) a. **Nonempty domains.** $A > C$ entails $A \diamondto C$.

b. **Stability.** C *cause* E entails $C > (C \textit{ cause } E)$.

c. **Idempotence.** $A \diamondto C$ entails $A > (A \diamondto C)$.

d. **Right weakening.**
If C entails C' then $A > C$ entails $A > C'$.

e. If C *cause* E is true, then C is not a subsentence of E.

Theorem

Given the assumptions in (4), Sartorio's Principle is equivalent to the Perfection Principle.

Proof ($\Rightarrow$)

Suppose Sartorio's Principle. Pick any sentences C and E and take $X = (C \text{ cause } E)$. Then by Stability, C *cause* E entails $C > X$. We also have the following chain of implications.

$$\begin{aligned}
C \textit{ cause } E \quad &\Rightarrow \quad \neg C > \neg(\neg C \textit{ cause } E) && \text{(Sartorio's Principle)} \\
&\Rightarrow \quad \neg(\neg C > (\neg C \textit{ cause } E)) && \text{(Nonempty domains)} \\
&\Rightarrow \quad \neg(C > (C \textit{ cause } E))[\neg C/C] && \\
& && (C \text{ is not a subsentence of } E) \\
&\Rightarrow \quad \neg(C > X)[\neg C/C] && (X = C \textit{ cause } E)
\end{aligned}$$

Hence C *cause* E entails $C > X$ and $\neg(C > X)[\neg C/C]$.

Proof ($\Leftarrow$)

Suppose the Perfection Principle. So $\neg C$ *cause* E entails $(C > X)[\neg C/C]$. Then by contraposition we have (†): $\neg(C > X)[\neg C/C]$ entails $\neg(\neg C$ *cause* $E)$. Observe the following chain of implications.

$$\begin{array}{rcll}
C \textit{ cause } E & \Rightarrow & \neg(C > X)[\neg C/C] & \text{(Perfection Principle)} \\
& \Rightarrow & \neg(\neg C > X[\neg C/C]) & \text{(Definition of } [\neg C/C]) \\
& \Rightarrow & \neg C \mathbin{\Diamond\!\!\rightarrow} \neg X[\neg C/C] & \text{(Definition of } \Diamond\!\!\rightarrow) \\
& \Rightarrow & \neg C > (\neg C \mathbin{\Diamond\!\!\rightarrow} \neg X[\neg C/C]) & \text{(Idempotence)} \\
& \Rightarrow & \neg C > \neg(\neg C > X[\neg C/C]) & \text{(Definition of } \Diamond\!\!\rightarrow) \\
& \Rightarrow & \neg C > \neg(C > X)[\neg C/C] & \text{(Definition of } [\neg C/C]) \\
& \Rightarrow & \neg C > \neg(\neg C \textit{ cause } E) & \text{(† and right weakening)}
\end{array}$$

- Presumably C *proto-cause* E entails $C > X$ or $\neg(C > X)[\neg C/C]$, for some X.

- Presumably C *proto-cause* E entails $C > X$ or $\neg(C > X)[\neg C/C]$, for some X.
- Our theorem gives us a way to turn *proto-cause* into *cause*.

- Presumably C *proto-cause* E entails $C > X$ or $\neg(C > X)[\neg C/C]$, for some X.
- Our theorem gives us a way to turn *proto-cause* into *cause*.
- **Case 1**. If C *proto-cause* E entails $C > X$ then add $\neg(C > X)[\neg C/C]$:

$$C \textit{ cause } E \qquad \text{if and only if } C \textit{ proto-cause } E \wedge \neg(C > X)[\neg C/C]$$

- Presumably C *proto-cause* E entails $C > X$ or $\neg(C > X)[\neg C/C]$, for some X.
- Our theorem gives us a way to turn *proto-cause* into *cause*.
- **Case 1**. If C *proto-cause* E entails $C > X$ then add $\neg(C > X)[\neg C/C]$:

$$C \textit{ cause } E \qquad \text{if and only if } C \textit{ proto-cause } E \wedge \neg(C > X)[\neg C/C]$$

- **Case 2**. If C *proto-cause* E entails $\neg(C > X)[\neg C/C]$ then add $C > X$.

$$C \textit{ cause } E \qquad \text{if and only if } C \textit{ proto-cause } E \wedge C > X$$

Plan

- Sufficiency for production

1. Difference making
   - Sartorio's analysis of difference-making
   - Unravelling Sartorio's Principle
   - The ubiquity of the Perfection Principle
   - On the pragmatic origins of the Perfection Principle

- Lewis (1973, p. 536) proposes that an event e causally depends on an event c just in case the following two counterfactuals are true: if c had occurred, e would have occurred.
- Wright (1985, 2011) proposes the NESS (Necessary Element of a Sufficient Set) test for causation, according to which something is a cause just in case there is a set of conditions that are jointly sufficient for the effect, but are not sufficient when the cause is removed from the set.

- Mackie's INUS condition states that a cause is "an insufficient but non-redundant part of a condition which is itself unnecessary but sufficient for the result" (Mackie 1974, p. 64).
- Beckers (2016) use a notion of *production*, arguing that the semantics of *is a cause of* involves comparing the presence and absence of the cause with respect to producing the effect. According to Beckers, C is an cause of E just in case, informally put, C produced E, and after intervening to make $\neg C$ true, $\neg C$ would not have also produced E.

Plan

- Sufficiency for production

1. Difference making
   - Sartorio's analysis of difference-making
   - Unravelling Sartorio's Principle
   - The ubiquity of the Perfection Principle
   - On the pragmatic origins of the Perfection Principle

Certainly, it seems to be the case that an inference can, historically, become part of semantic representation in the strict sense; thus, the development of the English conjunction since from a purely temporal word to a marker of causation can be interpreted as a change from a principle of invited inference associated with since (by virtue of its temporal meaning) to a piece of the semantic content of since.

(Geis and Zwicky 1971, pp. 565–566)

Iatridou (1993, 2021) observes that *then* in conditionals takes on a further meaning. She offers the following examples, which are unacceptable with *then* but fine without it.

(5) a. If I may be frank (*then) you are not looking good today.
b. If John is dead or alive (*then) Bill will find him.
c. If he were the last man on earth (*then) she wouldn't marry him.
d. Even if you give me a million dollars (*then) I will not sell you my piano.

Where $O(p)(q)$ denotes the conditional construction, Iatridou (1993) proposes that *if p then q* asserts $O(p)(q)$ and presupposes $\neg O(\neg p)(q)$.

$$C \ \wedge \ C \gg (C \ \textit{produce} \ E) \ \wedge \ \neg(\neg C \gg (\neg C \ \textit{produce} \ E))$$

(6) *C cause E* and *E because C* entail ...

a. Existential difference-making: $\neg(\neg C > (\neg C\ \textit{produce}\ E))$

b. Universal difference-making: $\neg C > \neg(\neg C\ \textit{produce}\ E)$

(7) a. Reyna was born at Royal Bolton Hospital but received a Danish passport because her mother was born in Copenhagen.

b. He has an American passport because he was born in Boston.

c. I think I was laid off because I'm 56 years old.

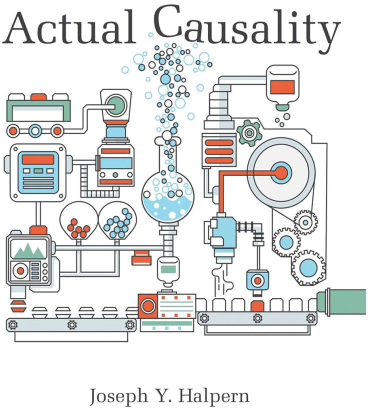
d. Naama Issachar ... could spend up to seven-and-a-half years in a Russian prison because 9.5 grams of cannabis were found in her possession during a routine security check.

e. A 90-day study in 8 adults found that supplementing a standard diet with 1.3 cups (100 grams) of fresh coconut daily caused significant weight loss.

(8) If he hadn't eaten 413 chicken nuggets, he wouldn't have been paralysed.

Summary

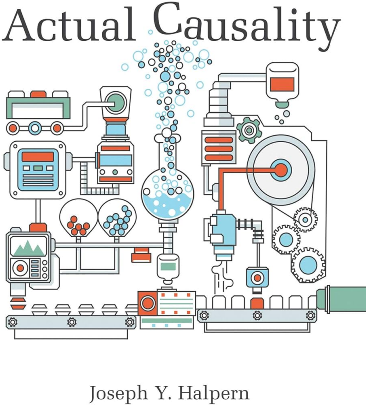
- Sartorio's Principle offers a principled way to distinguish the Billy and Suzy case from switching cases.
- But given the Principle's logical structure, one cannot simply add it to existing semantics of *cause*.
- Our theorem provides a way to add Sartorio's Principle to semantic theories of *cause*.



Halpern (2016), *Actual Causality*:

C is an actual cause of E just in case

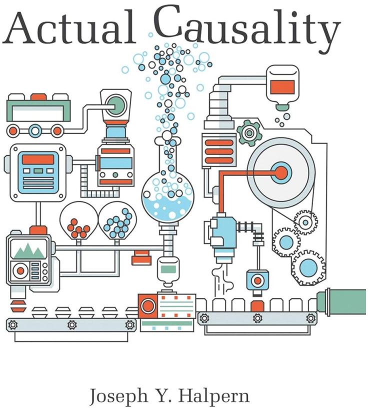
1. C and E actually occurred.



Halpern (2016), *Actual Causality*:

C is an actual cause of E just in case

1. C and E actually occurred.
2. There is a set of variables such that, holding them fixed at their actual values, if the cause had not occurred, the effect would not have occurred.



Halpern (2016), *Actual Causality*:

C is an actual cause of E just in case

1. C and E actually occurred.
2. There is a set of variables such that, holding them fixed at their actual values, if the cause had not occurred, the effect would not have occurred.
3. C is minimal: no proper subset of C satisfies (1) and (2).

$$
\begin{tikzcd}
ST \arrow[d] & & BT \arrow[d] \\
SH \arrow[rr] \arrow[dr] & & BH \arrow[dl] \\
& BS &
\end{tikzcd}
$$

$$
\begin{aligned}
SH &= ST \\
BH &= BT \wedge \neg SH \\
BS &= SH \vee BH
\end{aligned}
$$

Figure: Halpern's model of the Billy and Suzy case (2016, p. 31)

Halpern's account of the Billy and Suzy case

$$\begin{tikzcd}
ST = 1 \arrow[d] & & BT = 1 \arrow[d] \\
SH = 1 \arrow[rr] \arrow[dr] & & BH = 0 \arrow[dl] \\
& BS = 1 &
\end{tikzcd}$$

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

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

Halpern's account of the Billy and Suzy case

$$
\begin{tikzcd}
ST = 0 \arrow[d] & & BT = 1 \\
SH = 0 \arrow[dr] & & BH = 0 \arrow[dl] \\
& BS = 0 &
\end{tikzcd}
$$

$$SH = ST$$
$$BH = 0$$
$$BS = SH \vee BH$$

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

Two models of the switching scenario

$$\begin{tikzcd} S \arrow[d] \\ T \arrow[d] \\ A \end{tikzcd} \qquad T = \begin{cases} L & \text{if } S \\ R & \text{if } \neg S \end{cases} \qquad A = \begin{cases} 1 & \text{if } T = L \vee T = R \\ 0 & \text{otherwise} \end{cases}$$

(a) One-variable model

$$\begin{tikzcd} & S \arrow[dl] \arrow[dr] & \\ LT \arrow[dr] & & RT \arrow[dl] \\ & A & \end{tikzcd} \qquad \begin{aligned} LT &= S \\ RT &= \neg S \\ A &= LT \vee RT \end{aligned}$$

(b) Two-variable model

The two-variable model

$$
\begin{tikzcd}
 & S = 1 \arrow[dl] \arrow[dr] & \\
LT = 1 \arrow[dr] & & RT = 0 \arrow[dl] \\
 & A = 1 &
\end{tikzcd}
$$

$$
\begin{aligned}
LT &= S \\
RT &= \neg S \\
A &= LT \vee RT
\end{aligned}
$$

Figure: Two-variable model

The two-variable model

$$
\begin{tikzcd}
& S=0 \arrow[dl] & \\
LT=0 \arrow[dr] & & RT=0 \arrow[dl] \\
& A=0 &
\end{tikzcd}
$$

$$
\begin{aligned}
LT &= S \\
RT &= 0 \\
A &= LT \vee RT
\end{aligned}
$$

Figure: Two-variable model

Comparing the two models, Halpern and Pearl (2005, p. 872) write:

> *The two-variable model depicts the tracks as two independent mechanisms, thus allowing one track to be set (by action or mishap) to false (or true) without affecting the other. Specifically, this permits the disastrous mishap of flipping the switch while the left track is malfunctioning. More formally, it allows a setting where $S = 1$ and $RT = 0$. Such abnormal settings are imaginable and expressible in the two-variable model, but not in the one-variable model.*

The two-variable model also allows a setting where $S = 0$ and $RT = 0$. The one-variable model rules this out as part of its variable structure.

$$\begin{tikzcd} S \arrow[d] \\ T \arrow[d] \\ A \end{tikzcd} \qquad T = \begin{cases} L & \text{if } S \\ R & \text{if } \neg S \end{cases} \qquad A = \begin{cases} 1 & \text{if } T = L \vee T = R \\ 0 & \text{otherwise} \end{cases}$$

(a) One-variable model

$$\begin{tikzcd} & S \arrow[dl] \arrow[dr] & \\ LT \arrow[dr] & & RT \arrow[dl] \\ & A & \end{tikzcd} \qquad \begin{aligned} LT &= S \\ RT &= \neg S \\ A &= LT \vee RT \end{aligned}$$

(b) Two-variable model

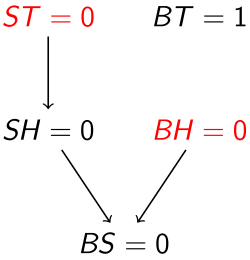
Figure: Two models of the switching scenario

In the two-variable model, one can intervene to make

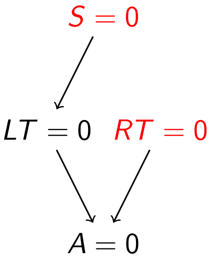
$$S = 0,\ LT = 0 \text{ and } RT = 0.$$

That is, interventions can make train disappear from the tracks!

The two-variable model



(a) Witness to Suzy causing the window to break



(b) Witness to the switch causing the train to arrive

The two-variable model

$$
\begin{tikzcd}
ST = 0 \arrow[d] & & BT = 1 \\
SH = 0 \arrow[dr] & & BH = 0 \arrow[dl] \\
& BS = 0 &
\end{tikzcd}
$$

(a) Witness to Suzy causing the window to break

$$
\begin{tikzcd}
& S = 0 \arrow[dl] & \\
LT = 0 \arrow[dr] & & RT = 0 \arrow[dl] \\
& A = 0 &
\end{tikzcd}
$$

(b) Witness to the switch causing the train to arrive

If Billy's rock can disappear mid-flight,
why can't the train disappear mid-journey as well?

Comparing the two models, Halpern and Pearl (2005, p. 872) write:

> *The two-variable model depicts the tracks as two independent mechanisms, thus allowing one track to be set (by action or mishap) to false (or true) without affecting the other. Specifically, this permits the disastrous mishap of flipping the switch while the left track is malfunctioning. More formally, it allows a setting where $S = 1$ and $RT = 0$. Such abnormal settings are imaginable and expressible in the two-variable model, but not in the one-variable model.*

Halpern's solution to the Billy and Suzy case is too sensitive to the choice of model.

Sartorio's Principle offers a more robust solution.

References I

- Beckers, Sander (2016). Actual Causation: Definitions and Principles. PhD thesis. KU Leuven. URL: https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1656621&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US.
- Geis, Michael L. and Arnold M. Zwicky (1971). On invited inferences. *Linguistic inquiry* 2.4, pp. 561–566. URL: www.jstor.org/stable/4177664.
- Hall, Ned (2000). Causation and the Price of Transitivity. *Journal of Philosophy* 97.4, pp. 198–222. DOI: 10.2307/2678390.
- Halpern, Joseph Y (2016). *Actual Causality*. MIT Press.
- Halpern, Joseph Y and Judea Pearl (2005). Causes and explanations: A structural-model approach. Part I: Causes. *The British journal for the philosophy of science* 56.4, pp. 843–887. DOI: 10.1093/bjps/axi147.
- Iatridou, Sabine (1993). On the contribution of conditional then. *Natural language semantics* 2.3, pp. 171–199.

References II

- Iatridou, Sabine (2021). Grammar matters. *Conditionals, Paradox, and Probability: Themes from the Philosophy of Dorothy Edgington*. Ed. by Lee Walters and John Hawthorne. Oxford University Press. DOI: 10.1093/oso/9780198712732.003.0008.
- Lewis, David (1973). Causation. *Journal of Philosophy* 70.17, pp. 556–567. DOI: 10.2307/2025310.
- Mackie, John L (1974). *The cement of the universe: A study of causation*. Clarendon Press. DOI: 10.1093/0198246420.001.0001.
- Sartorio, Carolina (2005). Causes As Difference-Makers. *Philosophical Studies* 123.1, pp. 71–96. DOI: 10.1007/s11098-004-5217-y.
- Wright, Richard (1985). Causation in tort law. *California Law Review* 73.6, pp. 1735–1828. DOI: 10.2307/3480373.
- — (2011). The NESS account of natural causation: a response to criticisms. *Perspectives on Causation*. Ed. by Richard Goldberg. Hart Publishing, pp. 13–66.