Causal models tutorial for COCOA

Bridget Copley

February 23, 2024

1 Introduction

1.1 What even is a causal model?

For our purposes, a causal model is a **formal representation** of the **structure** that **causal relations** give to our **conceptual model** of the world.

Basics (bedtime reading, sort of statisticky in parts): Pearl and Mackenzie (2018)

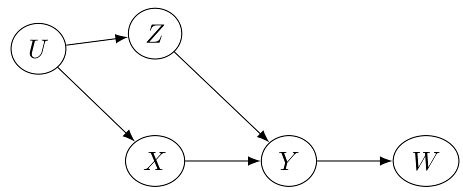
Basics (mathy): Pearl (2000); Halpern (2000); Halpern and Pearl (2005) ...

In linguistic semantics: Work since about 2005 by Stefan Kaufmann, Rebekah Baglini, Elitzur Bar-Asher Siegal, Cleo Condoravdi, Bridget Copley, Sven Lauer, Fabienne Martin, Prerna Nadathur, ...

- Causal structures are formally represented by means of a *directed acyclic graph* (DAG).
- There is a set V of variables that are the vertices (or nodes) of the graph.
- These are connected by a set of edges (or arrows) E.
- The edges are *directed* and represent the dependency of one value on another. For instance, $A \rightarrow B$ represents that the value of B is dependent in some way on the value of A.
- Absence of an edge between two variables means that the values are independent of each other.
- The dependencies are represented by functions.
- The causal model conveys that the values of some variables influence the values of other variables, according to both the arrows of the graph (which indicate the fact and direction of causation) and the functions associated with the structure (which give more information about which values go together).

- World knowledge provides information about which values go together, i.e. about which functions we are dealing with.

(1) U: season ({spring, summer, winter, fall})
X: sprinkler ({on, off})
Z: rain ({yes, no})
Y: wet ({yes, no})
W: slippery ({yes, no})



A plausible valuation for this model: U = summer, X = on, Z = no, Y = yes, W = yes.

U	X	Z	Y	W
summer	on	no	yes	yes
summer	off	no	no	no
summer	on	no	yes	yes
spring	off	yes	yes	yes
...	...	...	...	...

- Variables without arrows pointing at them are *exogenous* variables; their value depends only on circumstances that are not represented in the model (*background* variables, also *exogenous*). Variables with arrows pointing at them are *endogenous* variables.
- There is an asymmetry between exogenous variables and endogenous variables.
- Values of endogenous variables can be expressed by means of an equation over the variables that they depend on. When we do this, we can call the model a "structural equation model".
- The values are now going to be truth values, and we can express the nodes as *whether* propositions. If we want to relate this to lambda calculus, we can think of each node as being of type $\langle s, t \rangle$.
- The equation for the model in Table 1 is value(M?) = $\mathcal{F}_1(value(\text{L?}), value(\text{F?}))$

$$L? \longrightarrow F? \longleftarrow M?$$

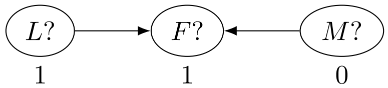
L? = whether there is lightning
F? = whether there is fire
M? = whether there is a lit match

$\mathcal{F}_1$:

L?	F?	M?
1	1	1
1	0	1
0	1	1
0	0	0

Figure 1: A collider with no actual values specified

A function is associated with different combinations of nodes. Each line of the table corresponds to a different situation. We can choose one of those situations to give an (actual) value the nodes in the structure, as in Figure 2:



L? = whether there is lightning

F? = whether there is fire

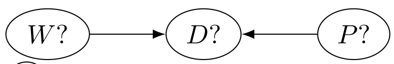
M? = whether there is a lit match

$\mathcal{F}_1$:

L?	M?	F?
1	1	1
1	0	1
0	1	1
0	0	0

Figure 2: The same collider with actual values specified

In the model in figures 1 and 2 above, the function $\mathcal{F}_1$ given by the table happens to be the *OR* function. In principle, we could set it to be any function we like that takes the values of the other variables and returns the value of F?, as long as it faithfully reflects world knowledge. See for example the model in figure 3, which has the same collider structure as the model shown above but which has a different function, $\mathcal{F}_2$.



W? = whether there is water in the tub

D? = whether water drains from the tub

P? = whether there is a plug in the drain

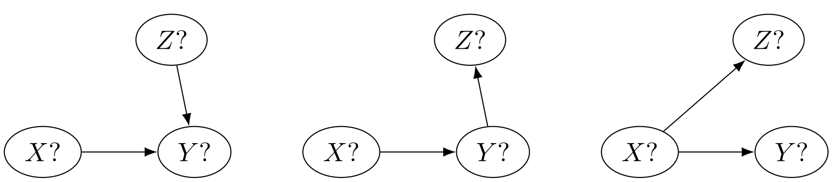
$\mathcal{F}_2$:

W?	D?	P?
1	1	0
1	0	1
0	1	0
0	0	0

Figure 3: Same structure, different function

(2) Taxonomy of basic 3-node structures:

Collider Chain Fork



1.2 FAQs

- How are the arrows to be read?
  - The arrows are NOT to be read as CAUSE, *leads to*, material implication, or Talmian forces.

– The arrows ARE to be read as
"the value of A affects the value of B",

"the value of A influences the value of B", or "the value of B listens to the value of A".

- Wait, don't causal models have to use probabilities and Bayes' Theorem?

  No.

- Which variable is THE cause?

  None is picked out. Some work on this: Bar-Asher Siegal and Boneh 2020.

- How does one decide which variables/arrows to include?

  – Short answer: We have intuitions about where the causal relations are supposed to be.

  – Long answer: We need to make sure that ...

    ... spurious dependencies do not get arrows (e.g. shoe size and reading ability)

    ... if A $\rightarrow$ B $\rightarrow$ C is in our model, the value of C only listens to the value of A through the value of B (e.g. A = number of cigarettes smoked daily, B average tar content of cigarettes smoked, C? = whether there is cancer)

    ... for any variable X in the structure, given X's immediate causal ancestors, X is independent of its (other) non-descendants (Causal Markov condition)

    ...

  – Linguistics answer: e.g. lexical and functional items give us causal relations

- But are we allowed to use a causal model to reason about causation? Is it not circular?

  No. Causation need not be derived from something else (pace David Lewis). It's ok to just point at a relation and say it's causal.

- This looks so... deterministic. Is that a problem?

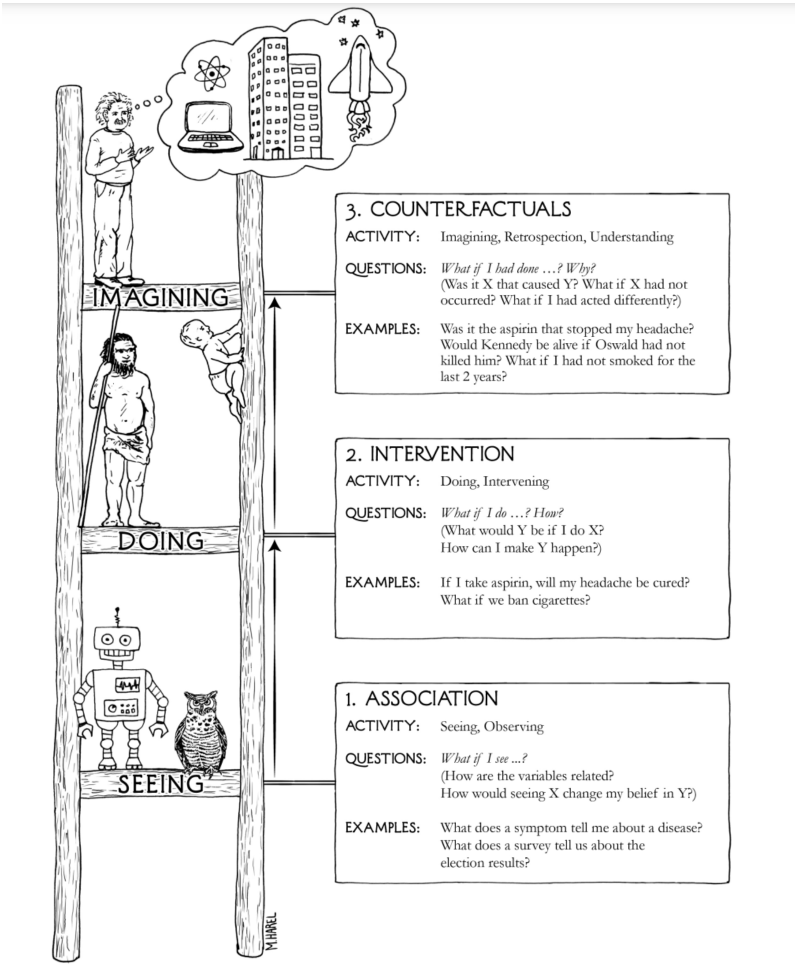
  Yes, but one that can be solved.

1.3 Where do causal models come from?

The short answer is that they come from the field of statistics in the early 20th century, with a huge ramp up of formalization and popularity in the last 30 years or so. See `https://plato.stanford.edu/entries/causal-models/`. Pearl and Mackenzie (2018) also has some history.

1.4 Why use causal models? Pearl's answer

Pearl and Mackenzie (2018) Chapter 1 (`http://bayes.cs.ucla.edu/WHY/why-ch1.pdf`) in particular gives some context:



(3) Level 2: Intervention (Pearl (2000), see also Woodward (2006))
$do(X = x)$: Set the value of X to x and erase all arrows that point to X

(4) Level 3: Counterfactuals

a. Pearl (2000) uses probabilities for these, e.g. in a case where $X \rightarrow Y$ is in the model, and we know that $X = 1$, we want to know the likelihood of $Y = 1$ if we (or something) had intervened to set X to 0. $P(Y_{X=0} \mid X = 1, Y = 1)$

1.5 Why use causal models? Linguistics answer

- We get the benefits of explicitly talking about causal relations, even when there are multiple influences and/or effects
- *Ceteris paribus*/closed-world condition: The existence of a causal relation doesn't entail the occurrence of the result, and no need to rule out irrelevant possibilities
- Interaction is easy to represent (e.g. progressives)
- Entrainment is easy to represent, bringing activities into alignment with causal treatments of accomplishments
- Better than mere dependency (e.g. Lewis 1973) in accounting for direct/indirect distinctions
- We can represent goal-directed action and control vs. accidental action
- We get the benefits of talking easily about causal powers (e.g. intentions, plans, authority, ability, dispositions)

2 How are causal models related to...

2.1 ...truth tables?

Traditional truth tables have Level 1 information only: **correlations** of the values, and both the **direction** (what values depend on what other values?) and **nature** (which relation/function is it?) of the dependency. The tables associated with causal models are similar, except that all of the above are assumed to be due to causality (kind of a Level 1.5?).

(5) Familiar truth tables:

And:

P	Q	$P \wedge Q$
1	1	1
1	0	0
0	1	0
0	0	0

Or:

P	Q	$P \vee Q$
1	1	1
1	0	1
0	1	1
0	0	0

Material conditional:

P	Q	$P \Rightarrow Q$
1	1	1
1	0	0
0	1	1
0	0	1

(6) Some causal relations, in (my, ad hoc) "truth table" notation:

Causal necessity	
A?	B?
1	{(0),1}
0	0

Causal sufficiency	
A?	B?
1	1
0	{(0),1}

2.2 ...event arguments?

Statisticians who use causal models talk about $X = x$ as an "event". We may or may not choose to represent a Davidsonian event with a node.

(7) a. I? $\rightarrow$ R?
I?: whether the subject has the intention that an R-result occurs
R?: whether the R-result occurs

b. I? $\rightarrow$ E? $\rightarrow$ R?
I?: whether the subject has the intention that an R-result occurs
E?: whether the subject is an agent of an E-event
R?: whether the R-result occurs

2.3 ...times?

(8) Two ways of looking at a timeline

a. The variables are relativized to times (Halpern and Pearl, 2005, 18): $X_{i_1}, X_{i_2}, X_{i_3}$ ...

b. Variables have values, and these values are relativized to times. The valuation function that could be redefined to take a time argument in addition to the variable argument. So, $value(X?)(i) = x$, or in a more semantics-friendly notation, $value(X?) = \lambda i.\ [\![\text{p}]\!](i)$.

2.4 ...possible worlds?

In general we can think of each line of the function ("truth table") as a possible world or situation.

Quantification over atomic possible worlds IS TO explicit causal relations

AS

Optimality Theory IS TO transformational phonology

3 Determinacy

As we asked earlier, isn't this kind of deterministic?

One solution: probabilities instead of truth values.

No need for probabilities if we use both Kratzer's causal premise semantics along with causal models. The premise background corresponding to the ordering source determines whether the value of Y is necessary or sufficient, given its parents, avoiding determinism.

A less complicated solution: Operations on nodes are permitted. The speaker expresses a commitment to a certain model at a certain time, perhaps putting it into a "Common Causal Model" as in Copley & Mari 2022, Copley & Kagan 2023. Uttered sentences operate dynamically on this model. Nodes and arrows can be added or subtracted from models if needed.

References

Halpern, J. Y. (2000). Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research 12*, 317–337.

Halpern, J. Y. and J. Pearl (2005). Causes and explanations: A structural-model approach. part i: Causes. *The British journal for the philosophy of science 56*(4), 843–887.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge University Press.

Pearl, J. and D. Mackenzie (2018). *The book of why: the new science of cause and effect*. Basic Books.

Woodward, J. (2006). Sensitive and insensitive causation. *The Philosophical Review 115*(1), 1–50.