



**HAL**  
open science

## From force dynamics to causal models

Bridget Copley

► **To cite this version:**

Bridget Copley. From force dynamics to causal models. Doctoral. Seoul, South Korea. 2021. hal-04455806

**HAL Id: hal-04455806**

**<https://hal.science/hal-04455806>**

Submitted on 13 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Seminar 2: From force dynamics to causal models

Bridget Copley, SFL (CNRS/Paris 8)

Seoul National University Linguistics Department  
21 October 2021

## 1 Benefits to having forces in the grammatical ontology

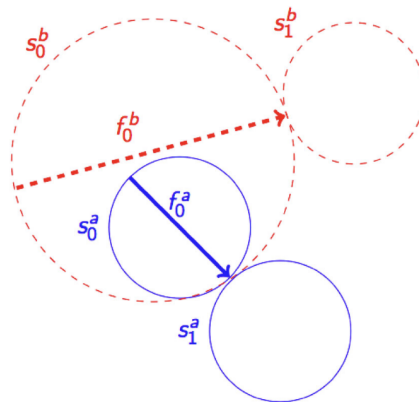
- We get the benefits of explicitly talking about causal relations

Quantification over primitive possible worlds IS TO explicit causal relations

AS

Optimality Theory IS TO transformational phonology

- *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
- Force interaction is easy to represent



- Entrainment is easy to represent, bringing activities into alignment with causal treatments of accomplishments
- Force-dynamic causation is better than mere dependency (e.g. Lewis 1973) in accounting for direct/indirect distinctions
- We can represent goal-directed action: the actor has an intention/goal in *s*, and from *s* arises the actor's action to bring about a situation of which the goal holds
- The Davidsonian argument is retained (though it is no longer a commonsense event)
- There is a natural distinction between events and states (namely, energy) that accounts for verbs of maintaining

## 2 Two issues with Copley & Harley's force-theoretic framework

### 2.1 Dynamic verbs don't always describe an input of energy

- (1)
  - a. The ball rolled down the hill.
  - b. The soup cooled.
  - c. John let the children eat.
  - d. Bill killed my plants by not remembering to water them.

### 2.2 Causal statives

- (2)
  - a. That wall supports the ceiling.
  - b. That wall must support the ceiling.
  - c. In case you want to know, that wall supports the ceiling.
- (3)
  - a. That curtain lets the light through.
  - b. That curtain must let the light through.
  - c. In case you want to know, that curtain lets the light through.
- (4)
  - a. The flowers decorate the room beautifully.
  - b. The flowers must decorate the room beautifully.
  - c. In case you want to know, the flowers decorate the room beautifully.
- (5)
  - a. #John decorates the room beautifully.
  - b. John must decorate the room beautifully.
  - c. In case you want to know, John decorates the room beautifully.

### 3 Causal models

#### 3.1 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.

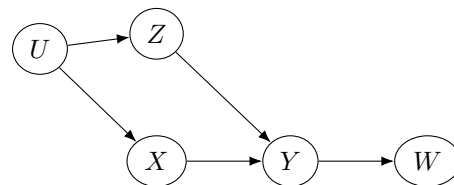
Basics (bedtime reading, sort of statisticky in parts): Pearl and Mackenzie (2018)

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

In linguistic semantics: Work by Stefan Kaufmann in the early 2000s, and since about 2010 by Rebekah Baglini, Elitzur Bar-Asher Siegal, Cleo Condoravdi, Sven Lauer, Fabienne Martin, Perna 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,  $\textcircled{A} \rightarrow \textcircled{B}$  represents that the value of  $\textcircled{B}$  is dependent in some way on the value of  $\textcircled{A}$ .
- Absence of an edge between two variables means that the values are independent of each other.

- (6)  $U$ : 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:  $U = \text{summer}$ ,  $X = \text{on}$ ,  $Z = \text{no}$ ,  $Y = \text{yes}$ ,  $W = \text{yes}$ .

$U$	$X$	$Z$	$Y$	$W$
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”.

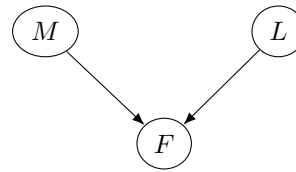
(7) An example from Halpern and Pearl (2005) with truth values

$F = 1$  if *There is fire* is true,  $F = 0$  if it is false. (“whether there is fire”)

$M = 1$  if *The match is lit* is true,  $M = 0$  if it is false. (“whether the match is lit”)

$L = 1$  if *There is lightning* is true,  $L = 0$  if it is false. (“whether there is lightning”)

Function that takes  $L$ ,  $M$  and yields value of  $F$ :  $F = 1$  if either  $L = 1$  or  $M = 1$ .



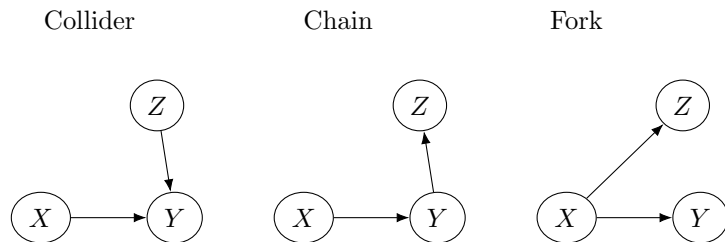
### 3.2 A formal setup

From Halpern and Pearl (2005):

- (8) A *signature*  $\mathcal{S}$  is a tuple  $(\mathcal{U}, \mathcal{V}, \mathcal{R})$  where  $\mathcal{U}$  is a set of endogenous variables,  $\mathcal{V}$  is a set of exogenous variables, and  $\mathcal{R}$  associates with every variable  $Y \in \mathcal{U} \cup \mathcal{V}$  a nonempty set  $\mathcal{R}(Y)$  of possible values for  $Y$ .
- (9) A *causal model* (or *structural model*) is a tuple  $M = (\mathcal{S}, \mathcal{F})$ , where  $\mathcal{F}$  associates with each variable  $X \in \mathcal{V}$  a function denoted  $F_X$  such that:  $F_X = (\times_{U \in \mathcal{U}} \mathcal{R}(U)) \times (\times_{Y \in \mathcal{V} - X} \mathcal{R}(Y)) \rightarrow \mathcal{R}(X)$
- In words:  $F_X$  is a function that determines the value of  $X$  given the values of all the other variables.
  - Notation: The set of all ordered pairs  $(a, b)$ , where  $a$  is an element of  $A$  and  $b$  is an element of  $B$ , is called the Cartesian product of  $A$  and  $B$  and is denoted by  $A \times B$ .
  - More notation:  $(\times_{U \in \mathcal{U}} \mathcal{R}(U))$  is the ordered  $n$ -tuple of all the values of all the variables in  $\mathcal{U}$ .
  - Exercise: How might we describe  $F_X$  in words, using “the arrows that...”?
- (10) Example: if  $F_X(Y, Z, U) = Y + U$ , then if  $Y = 3$  and  $U = 2$ , then  $X = 5$  no matter how

$Z$  is set. (As shorthand, we can also write this equation as:  $X = Y + U$ ) The function  $\mathcal{F}$  defines a set of these, called *structural equations*.

(11) Taxonomy of basic 3-node structures:



### 3.3 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  $\textcircled{A}$  affects the value of  $\textcircled{B}$ ”, “the value of  $\textcircled{A}$  influences the value of  $\textcircled{B}$ ”, or “the value of  $\textcircled{B}$  listens to the value of  $\textcircled{A}$ ”.

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

- No, they can use any value.

Which variable is THE cause?

- See Baglini & Bar-Asher Siegal 2020

How does one decide which variables to include?

- Short answer: Include everything that’s relevant.
- Long answer: See next long answer.

How does one decide which arrows to include?

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

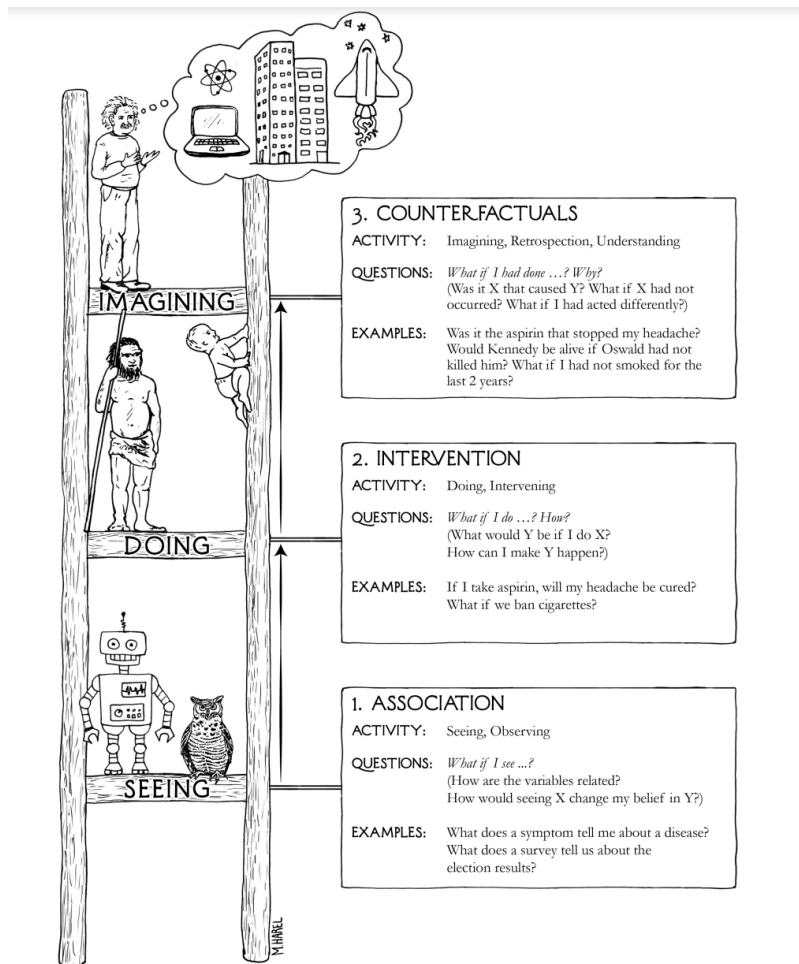
- Long answer: We need to make sure that ...
  - ... spurious dependencies do not get arrows (e.g. shoe size and reading ability)
  - ... if  $\textcircled{A} \rightarrow \textcircled{B} \rightarrow \textcircled{C}$  is in our model, the value of  $\textcircled{C}$  only listens to the value of  $\textcircled{A}$  through the value of  $\textcircled{B}$
  - ... for any variable  $\textcircled{X}$  in the structure, given  $\textcircled{X}$ 's immediate causal ancestors,  $\textcircled{X}$  is independent of its (other) non-descendants (Causal Markov condition)
  - ...

### 3.4 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.

### 3.5 Why use causal models?

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



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

- (13) Level 3: Counterfactuals
- Pearl (2000) uses probabilities for these, e.g. in a case where  $\mathbb{X} \rightarrow \mathbb{Y}$  is in the model, and we know that  $\mathbb{X} = 1$ , we want to know the likelihood of  $\mathbb{Y} = 1$  if we (or something) had intervened to set  $\mathbb{X}$  to 1.  $P(\mathbb{Y}_{\mathbb{X}=0} = 1 \mid \mathbb{X} = 1, \mathbb{Y} = 1)$
  - Kaufmann (2013): 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  $\mathbb{Y}$  is necessary or sufficient, given its parents, avoiding determinism.



## 4 How are causal models related to...

### 4.1 ...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?).

(14) Familiar truth tables:

And:			Or:			Material conditional:		
$P$	$Q$	$P \wedge Q$	$P$	$Q$	$P \vee Q$	$P$	$Q$	$P \Rightarrow Q$
1	1	1	1	1	1	1	1	1
1	0	0	1	0	1	1	0	0
0	1	0	0	1	1	0	1	1
0	0	0	0	0	0	0	0	1

(15) Some relations and functions to think about:

Causal necessity

Ⓐ	Ⓑ
1	{(0,1)}
0	0

Causal sufficiency

Ⓐ	Ⓑ
1	1
0	{0,(1)}

Stimulatory influence (+):

Ⓐ	Ⓑ
1	1
0	0

Inhibitory influence (-):

Ⓐ	Ⓑ
1	0
0	1

Ⓐ	Ⓑ
1	.1
2	.01
3	.001
4	.0001

Ⓐ	Ⓑ
.0006	55,679
.0008	16,450
.0010	9611
.0012	5001

## 4.2 ...force interactions?

Force interactions are easy to represent as colliders. Two ways to figure out the value of a node that has incoming arrows:

- Value of a node is determined from the values of all its influences (standard causal model)
- Each arrow represents a function and incoming arrows “compete” to determine the value of the node (more like force dynamics)

## 4.3 ...intentions?

If one is doing an action intentionally, the value of the result listens to the value of the intention. So the causing node could be an intention.

## 4.4 ...times/situations?

The variables can be relativized to times (Halpern and Pearl, 2005, 18):  $X_{i_1}, X_{i_2}, X_{i_3} \dots$

- (16)
- a. Launching (for telicity): time of evaluation of influencing node < time of evaluation of affected node
  - b. Entrainment (for atelicity): time of evaluation of influencing node = time of evaluation of affected node

Another possibility: Variables have values, and these values are relativized to times. The valuation function (call it  $\mathcal{R}$ ) could be redefined to take a time argument in addition to the variable argument. So,  $\mathcal{R}(X)(i) = x$ , or in a more semantics-friendly notation,  $\mathcal{R}(X) = \lambda i. \llbracket p \rrbracket(i)$ .

## 4.5 ...event arguments?

Statisticians who use causal models talk about  $X = x$  as an “event”. We can represent a Davidsonian event as a node.

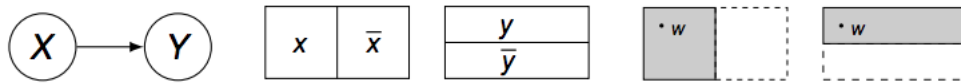
- (17)
- a. John was buttering the toast in the kitchen.
  - b. John was slowly buttering the toast.
- (18)
- a. \*In the hall, John was buttering the toast in the kitchen.
  - b. \*Quickly, John was slowly buttering the toast.

Nominalizations (Data from Sichel 2010, Alexiadou et al 2013)

- (19) a. The scientists justified the evacuation.  
 $\textcircled{P} \rightarrow \textcircled{E}$   
 $\textcircled{P}$  = whether  $\exists p : p(\textit{scientists})(i)$   
 $\textcircled{E}$  = whether  $\textit{occur}(e_{\textit{justification-of-evacuation}})(i')$
- b. The hurricane justified the evacuation.  
 $\textcircled{P} \rightarrow \textcircled{J}$   
 $\textcircled{P}$  = whether  $\exists p : p(\textit{hurricane})(i)$   
 $\textcircled{J}$  = whether  $\textit{justified}_{\textit{adj}}(x_{\textit{evacuation}})(i')$
- c. The hurricane destroyed the marina.  
 $\textcircled{P} \rightarrow \textcircled{E}$   
 $\textcircled{P}$  = whether  $\exists p : p(\textit{hurricane})(i)$   
 $\textcircled{E}$  = whether  $\textit{occur}(e_{\textit{destruction-of-the-marina}})(i')$
- (20) a. The scientists' justification of the evacuation  
*te...*
- b. \*The hurricane's justification of the evacuation  
 No event argument to bind!
- c. The hurricane's destruction of the marina  
*te...*

#### 4.6 ...possible worlds?

Kaufmann (2013):



Causal Networks	Possible worlds
Outcome	Possible world
Event	Proposition
Variable	Partition

From there we have a choice: What do we need from the familiar Kratzer-style tools (atomic possible worlds, sets of atomic possible worlds, orderings on sets of atomic possible worlds)? And

if we don't use them, do we need to use probabilities?

## 5 What can causal models do for semantics?

Many of the same things as the force-theoretic model:

- 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
- 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 (intention node)
- The Davidsonian argument is can be retained (~~though it is no longer a commonsense event~~)
- ~~There is a natural distinction between events and states (namely, energy)~~

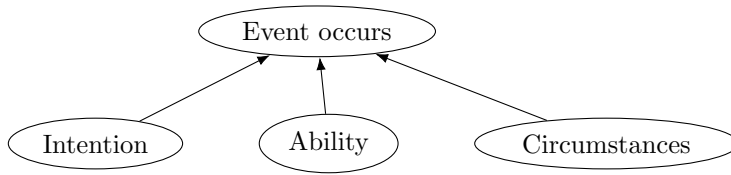
Dynamic verbs don't always describe an input of energy, and that's fine:

- (21)
- a. The ball rolled down the hill.
  - b. The soup cooled.
  - c. John let the children eat.
  - d. Bill killed my plants by not remembering to water them.

Causal statives: energy is detached from causation, no problem representing causation in the absence of energy

Some more vague but still useful reasons to take up causal models:

- More clarity around what the arrows are: they represent dependencies between values
- Influences are much easier to individuate
- More freedom with functions/relations (too much though?)



## 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.
- Kaufmann, S. (2013). Causal premise semantics. *Cognitive science* 37(6), 1136–1170.
- 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.