Causal necessity, causal sufficiency,

and the implications of causative verbs.

Sven Lauer (University of Konstanz)
Joint work with Prerna Nadathur (Stanford)

Modeling causal dependencies in formal semantics
University of Konstanz
May 22, 2017

The Basic puzzle

What does causative *make* mean?

(1) a. John made the children dance.
    b. John caused the children to dance.

- Both sentences: John ‘brought about’ the dancing of the children.
- (1-a) says more.
- But what?

- **Preliminary answer:** that the children did not have a say in the matter, that the will of the children was immaterial.
- This is what I call the coercive implication of *make*.

The Basic Puzzle (cont'd)

What does causative *make* mean?

(2) $make(S,O,P) = cause(S,O,P)$ & if O had not wanted $P(O)$ to come about, it still would have.

- Done.
- Not so fast!
- *Make* can take non-volitional, even inanimate surface objects (='causees').

  (3) The sun made the flowers wilt.

Problem

- We want to predict the coercive implication.
- We cannot make reference to the causee's volitional state.

The Basic Puzzle (summary)

What does causative *make* mean?

Wanted:

- A **unified** semantics for *make*, . . .
- that predicts the coercive implication with volitional causees, . . .
- but also applies to non-volitional causees.

The Plan

1. The claim: *make* is not a hyponym of *cause*
2. The intuition: *make* predicates causal sufficiency
3. The first attempt: Causal sufficiency in causal models
4. The second attempt: Into the wilderness . . .

The claim: *make* is not a hyponym of *cause*

A misleading intuition

make is not a hyponym of *cause*

(2) $make(S, O, P) = cause(S, O, P)$ & ...

- Plausible, widely shared intuition (Lewis 1973, and many after him):
  - *cause* entails ‘counterfactual necessity’
  - i.e. (4) entails (5):

(4) The recession caused John to lose his house.

(5) (Other things being equal,) If the recession had not happened, John would not have lost his house.

A misleading intuition

make is not a hyponym of *cause*

Claim: *make* does **not** entail counterfactual necessity.

- i.e., despite appearances, (6) does not entail (7):

(6) Society made me kill.

(7) If society had not been the way it was, I would not have killed.

Example (Failure of necessity-entailment)

(8) Last year, I was not sure if I should go to band camp, but then my mother insisted that I go. I am so happy she made me go: I had the best summer ever.

Does not entail: If the mother of the speaker had not insisted, the speaker would not have gone to band camp.

A misleading intuition

make is not a hyponym of *cause*

MADDOW: You worked for [the health insurance company] CIGNA for 15 years, you left last year. What **caused** you to change your mind about what you were doing and to leave?

POTTER: Well, two things. One, it was kind of gradually. One instance or in one regard because I was becoming increasingly skeptical of the kinds of insurance policies that the big insurance companies are promoting and marketing these days. [. . .]

The other thing that really **made** me make this final decision to leave the industry occurred when I was visiting family in Tennessee a couple of summers ago, and [narrates the experience of happening on a ‘healthcare expedition’ where uninsured patients were treated by volunteer doctors in animal stalls at a fairground].

make is not a hyponym of *cause*

Summary

- *make* does not entail counterfactual necessity.
- New questions:
  - How do we characterize the ‘bringing about’ component of *make*?
  - Why does it often **seem** as if *make* entailed counterfactual necessity?

The intuition: *make* predicates causal sufficiency

Introducing sufficiency

- Idea: *make* predicates Sufficiency (as opposed to necessity)
- Let us assume that (9), roughly, means (10):

(9) The thunderstorm made the children scream.

(10) The thunderstorm was sufficient for the screaming of the children.

- What does (10) amount to?
- Preliminary characterization: An event being sufficient for another means that the first **ensures** that the second happens.

(11) The thunderstorm ensured that the children screamed.

(12) Given the thunderstorm, the children could not but scream.

Introducing sufficiency

Benefit 1: The coercive implication comes for free

Assumption:

A made B VP $\simeq$ A ensured that B VPed.

- We have not encoded the coercive implication directly.
- Yet, the hope is that it arises indirectly **when the embedded eventuality is a volitional action**.

(1a) John made the children dance.
$\Rightarrow$ What John did ensured that the children would dance.

- Given sufficiency, the children cannot have acted freely.
  - Suppose they did.
  - Then they could have decided to do otherwise.
  - But then, it is not appropriate to say that John **ensured** that the children would dance.

Introducing sufficiency

Benefit 2: Causal Perfection

Assumption:

A made B VP $\simeq$ A was sufficient for B's VPing.

- There is a strong tendency to interpret statements of sufficiency as asserting necessity, as well.
- This is known as conditional perfection Geis and Zwicky (1971), van der Auwera (1997), Horn (2000), von Fintel (2001), Franke (2009)

(13) If you study for the exam, you will get an A.
$\rightsquigarrow$ If you don't study for the exam, you will not get an A.

(14) If you study for the exam, you will get an A. Actually, even if you don't study you might get an A.

Introducing sufficiency

Benefit 2: Causal Perfection

Claim

make predicates sufficiency, the necessity implications come about pragmatically through perfection.

- Challenge: There should be contexts where the necessity can be coherently denied.
- And there are:

(15) My husband's arrest (finally) made me get a divorce. . . . Even if his arrest had not made me do it, I might have gotten a divorce anyways, given the way he treated me.

Introducing sufficiency (Summary)

Summary

Assuming that *make* predicates sufficiency ...

- ...allows to capture the coercive implication without hard-coding it.
- ...gives us a handle on the perceived necessity implications of *make*, explaining them as instances of perfection.

The first attempt: Causal sufficiency in causal models

Definition

Dynamics

A *dynamics* is a triple $\langle P, B, F \rangle$ where

- P is a set of binary variables,
- $B \subseteq P$ is a distinguished set of background variables,
- F is a function that assigns each non-background variable:
  - an n-tuple Z_X of variables in P,
  - a function $f_X : \{0,1\}^n \rightarrow \{0,1\}$

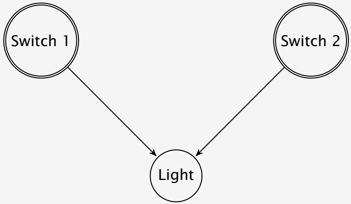
  F is 'rooted' in B.

  (causal chains acyclic, every chain starts in B)

Dynamics

Lifschitz Example

"Suppose there is a circuit such that the light is on (L) exactly when both switches are in the same position (up or not up)."



$f_{\text{Light}} =$

Switch 1	Switch 2	Light
0	0	1
0	1	0
1	0	0
1	1	1

Particular facts

Worlds and situations

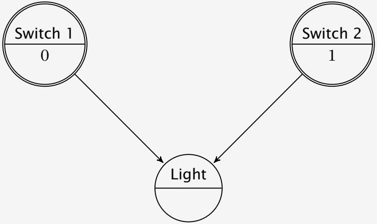
- A world is total function $w : P \rightarrow \{0, 1\}$
- A situation is total function $s : P \rightarrow \{0, 1, u\}$.
  - u stands for 'undefined'.
  - *i.e.*, situations are partial specifications of worlds.

"Suppose there is a circuit such that the light is on (L) exactly when both switches are in the same position (up or not up). At the moment switch 1 is down ($\neg S_1$), switch 2 is up (S_2)."

s :

S_1	0
S_2	1
L	u

Lifschitz example



$f_{\text{Light}} =$

Switch 1	Switch 2	Light
0	0	1
0	1	0
1	0	0
1	1	1

Running the dynamics

The function $\mathcal{T}$

The function $\mathcal{T}_D$

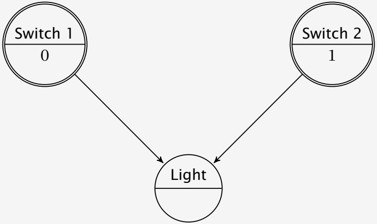
For $D = \langle P, B, F \rangle$ a dynamics, let $\mathcal{T}_D$ be the following function from situations to situations: For any situation s and $q \in P$:

- If $q \in B$, then $\mathcal{T}_D(s)(q) = s(q)$. background variables unchanged
- Otherwise,
  - If $s(q) \neq u$, then $\mathcal{T}_D(s)(q) = s(q)$. variables retain their values
  - If $s(q) = u$ and $Z_q = \langle p_1, \ldots, p_n \rangle$ and $f_q(s(p_1), \ldots, s(p_n)))$ defined, then $\mathcal{T}_D(s)(q) = f(s(p_1, \ldots, p_n))$.
    undefined variables set per F, if possible
  - Otherwise, $\mathcal{T}_D(s)(q) = u$.

- $\mathcal{T}_D$ takes a situation and 'runs the dynamics for one step'.
- Variables that are already set are maintained.
- Unset variables get the value that F specifies (if any).

Lifschitz example

Situation s:

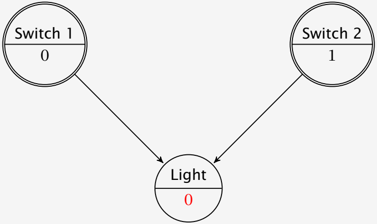


$f_{\text{Light}} =$

Switch 1	Switch 2	Light
0	0	1
0	1	0
1	0	0
1	1	1

Lifschitz example

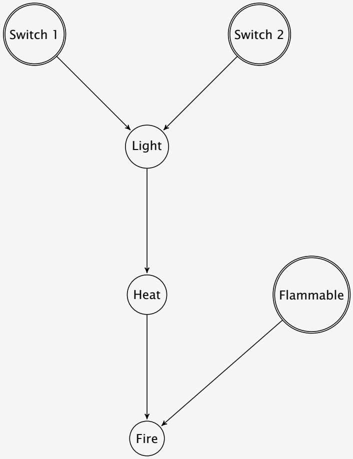
Situation $\mathcal{T}_D(s)$:



$f_{\text{Light}} =$

Switch 1	Switch 2	Light
0	0	1
0	1	0
1	0	0
1	1	1

Extended Lifschitz example



- f_{Light} as before.
- $f_{\text{Heat}} =$

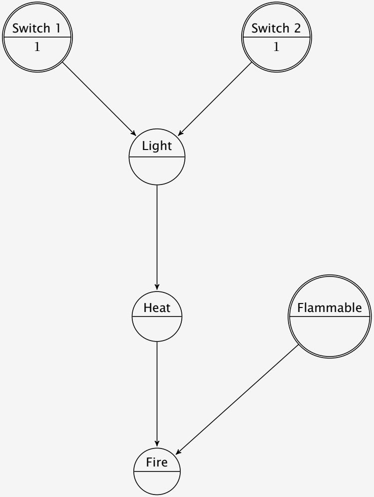
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example

Situation s:



- f_{Light} as before.
- $f_{\text{Heat}} =$

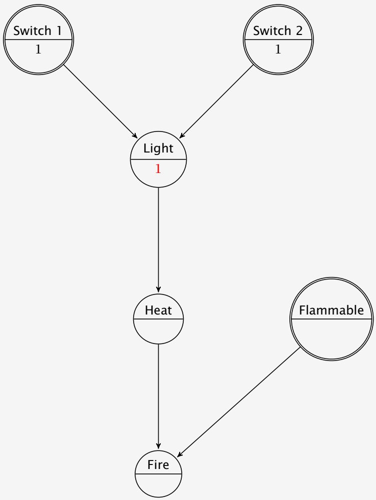
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example

Situation $\mathcal{T}_D(s)$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

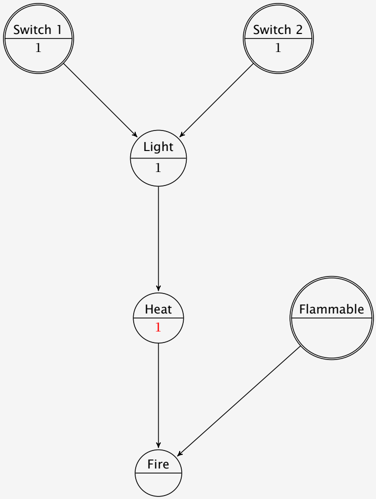
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example

Situation $\mathcal{T}_D(\mathcal{T}_D(s))$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Causal support

- Given ‘rootedness in B’ and finite P, T_D has a finite fixed point for any s.
- For any s, let s^* denote the least fixed point of $\mathcal{T}_D$.

Causal support

Let D a dynamics and s a situation for D. Then for any ϕ:

$$s \models_D \phi \text{ iff } s^* \vDash \phi$$

(ϕ a formula of propositional language with propositional letters P, $\vDash$ its classical semantics)

Causal sufficiency

First attempt

Natural:

- A situation is causally sufficient for ϕ iff

$$s \models_D \phi$$

(Baglini and Francez 2015, Nadathur 2016)

- But we need a notion of sufficiency suitable for facts (or formulas), relative to a situation.

Causal sufficiency

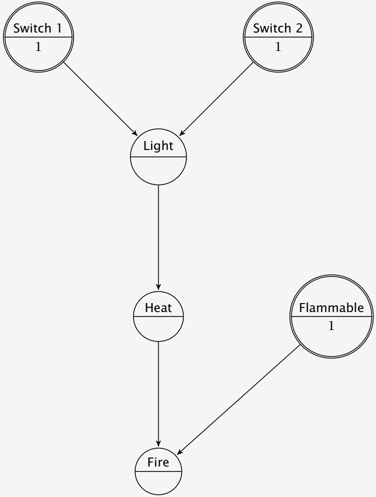
For $p \in P$ and situation s:

$$p \rightsquigarrow_s \phi \text{ iff } s \cup \{\langle p, 1 \rangle\} \models_D \phi$$

A fact $X = x$ is sufficient for ϕ, relative to D and s, if:
s augmented with $X = x$ causally entails ϕ.

Extended Lifschitz example

$Flammable \rightsquigarrow_s$ Fire because $s[Flammable/1]$ is:



- f_{Light} as before.
- $f_{\text{Heat}} =$

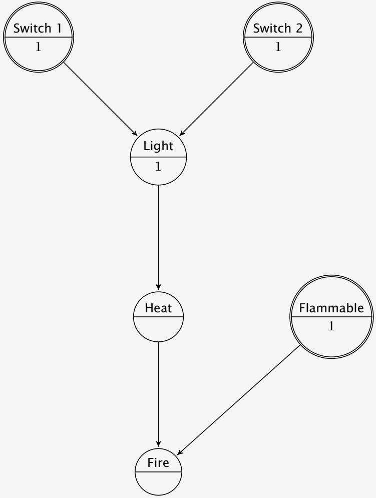
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example

$Flammable \rightsquigarrow_s$ Fire because $\mathcal{T}_D(s[Flammable/1])$ is:



- f_{Light} as before.
- $f_{\text{Heat}} =$

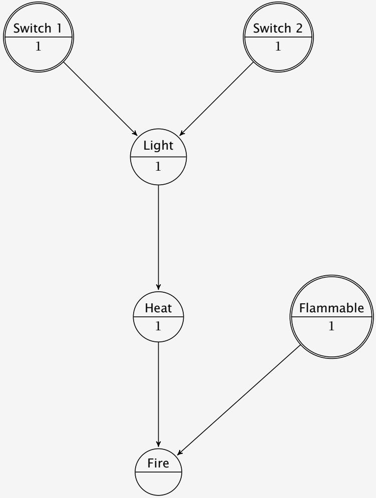
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example

$Flammable \rightsquigarrow_s$ Fire because $\mathcal{T}_D(\mathcal{T}_D(s[Flammable/1]))$ is:



- f_{Light} as before.
- $f_{\text{Heat}} =$

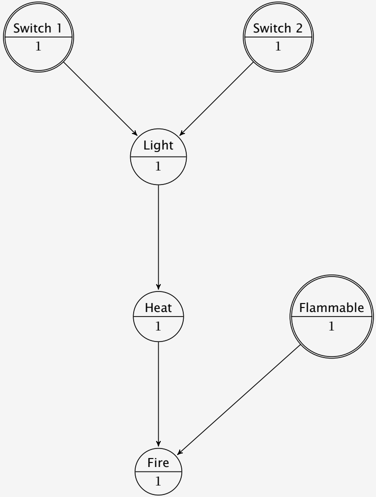
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example

$\mathit{Flammable} \rightsquigarrow_s$ Fire because
$\mathcal{T}_D(\mathcal{T}_D(\mathcal{T}_D(s[\mathit{Flammable}/1]))) = s[\text{Flammable}/1]^*$ is:



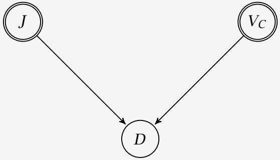
- f_{Light} as before.
- $f_{\text{Heat}} =$

Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

A problem

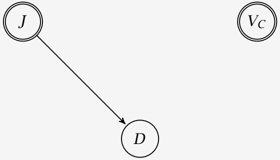


$$f_D = \begin{array}{cc|c} J & V_C & D \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{array}$$

- J: some action of John
- V_C: whether the children wanted to dance
- D: whether the children dance

Given the dynamics we want to say $J \rightsquigarrow_{\varnothing} D$, but this is not the case.

A problem



$f_D =$

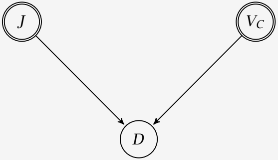
J	D
0	0
1	1

- J: some action of John
- V_C: whether the children wanted to dance
- D: whether the children dance

Now it is true that $J \rightsquigarrow_{\varnothing} D$. That is, relative to the empty situation, what John did was causally sufficient for the dancing of the chidlren. Success?

Another problem

More serious.



$f_D =$

J	V_C	D
0	0	0
1	0	0
0	1	0
1	1	1

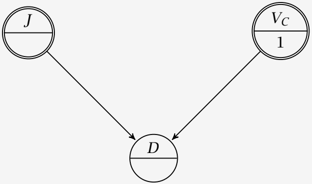
- J: some action of John
- V_C: whether the children wanted to dance
- D: whether the children dance

Given the dynamics we **do not want to say** that what John did was sufficient for the dancing of the children. Maybe we even want to say that it **could not have been sufficient.**

Another problem

More serious.

Consider a situation s' such that $s'(V_C) = 1$:



$f_D =$

J	V_C	D
0	0	0
1	0	0
0	1	0
1	1	1

- J: some action of John
- V_C: whether the children wanted to dance
- D: whether the children dance

Here we have $J \rightsquigarrow_{s'} D$. This may be fine for a notion of sufficiency in general, but not for *make*: If the children happened to want to dance, we would not count *John made the children dance* as true in this situation.

The problem

- If we want the sufficiency implication of *make* to predict the coercive implication, the current definition of sufficiency is not appropriate.
- We would somehow have to ensure that, whenever the embedded VP denotes an action then:
  (i) There is a variable present that represents whether the caussee wanted to do the action. arguably unproblematic
  (ii) This variable is **not** set by the situation relative to which we judge sufficiency. not clear how to ensure this

The second attempt: Into the wilderness . . .

Diagnosis

- Guess: The problem is that we do not quite capture the initial intuition.
- Because we cannot talk about time.
- Suppose that the children happened to want to dance.
- Idea: Then (16) appears to still imply that the children did not make a free decision to dance, because even if they had changed their mind after what John did, they still would have danced.
  - Or: They could not have changed their minds after what John did. More on that later.

(16) John made the children dance.

Bringing in time

- I assume a moment-based system of branching time (cf. Antje's talk?).
- We start with a set of *moments*, which are ordered into a tree-like structure by an order $<$.
- I am going to assume discrete time-steps, every moment has a unique latest predecessor.
- Maximal $<$-chains of moments are histories.

Causal dynamics in time

- The interpretation of the variables in the causal dynamics changes:
  - They no longer stand for propositions, but rather for fluents (think: predicates of times).
  - Their value can change over times.
- Moments are situations in the sense defined above, i.e. partial interpretation functions for the variables.
- Idea:
  - We define what it is for a history to miracle-free after a moment m.
  - This is the case if the non-background variables only change according to the dynamics.
  - Background variables can change at any time step, without a miracle.
  - Sufficiency of p at moment m is then determined by quantification over all miracle-free-after-m histories.

Rootedness and $\mathcal{T}_D$

- In Schulz (2007, 2011), dynamics are required to rooted, which implies that the 'causal ancestor'-relation is acyclic.
- And the $\mathcal{T}_D$-function retains all previously set values.
- This is what ensures that the $\mathcal{T}_D$ function has finite fixed-points.
- Arguably, using $\mathcal{T}_D$ for specifying miracle-free-ness and requiring rootedness have counterintuitive consequences in the present setting.
  - *e.g.*, it would be guaranteed that any mircale-free history falls into a stable state after finite time.
  - And, arguably, we want to allow for cycles in the 'causal ancestor' relation.

Rootedness and $\mathcal{T}_D$

An illustration

https://www.youtube.com/watch?v=CWqoI7UKpQg

Rootedness and $\mathcal{T}_D$

- So I propose to give up rootedness.
  - Do we need something weaker in its place?
- And I define an alternative to $\mathcal{T}_D$, in two steps:

Causal effects

Let m be a moment. Then $S_D(m)$ is a situation such that for all $q \in P$:

- If $Z_q = \langle p_1, \ldots, p_n \rangle$ and $f_q(m(p_1), \ldots, m(p_n)))$ defined, then $S_D(m)(q) = f(m(p_1, \ldots, m(p_n)))$.
- Otherwise, $S_D(m)(q) = u$.

S_D sets all nnon-background variables that can be set per F on the basis of facts from m, leaves all others undefined.

Inertia

Inertial completion

In is that function from pairs $\langle s, m \rangle$ to situations such that for all $m, s, q \in P$:

- If $q \notin B$ and $s(q) = u$ then $In(s, m)(q) = m(q)$.
- Otherwise, $In(s, m)(q) = s(q)$.

The inertial completion of s with m 'copies' the non-background variables from m if they are 'undefined' in s.

Miracle-free-ness

Miracle-free after m

A history h with $m \in h$ is miracle-free after m iff for all $m' \in h$ such that $m < m'$ (with m'^* the immediate predecessor of m'):

$$m' \upharpoonright (P - B) = In(S_D(m'^*), m'^*) \upharpoonright (B - P)$$

- Idea: h is miracle-free after m iff all moments m' after m are derived from their immediate predecessor, by:
  - running the causal dynamics one step;
  - performing inertial completion on the rest;
  - setting the background variables arbitrarily.

Causative *make* in time

Causative *make*: an attempt

$make(p, \phi)$ is true at m iff:

- $m(p) = 1$
- ϕ is true eventually all histories h that are miracle-free after m
- ϕ is **not** true eventually in all histories h' that are miracle-free after $m' = m^{[p \mapsto 0]}$

Then (I would have hoped):

- If there is a background variable X representing the will of the agent to make ϕ true, it is guaranteed that ϕ does not causally depend on the value of X.
- If there is a non-background variable X representing the will of the agent to make ϕ true, it is guaranteed that
  - either ϕ does not causally depend on the value of X,
  - if ϕ requires $X = 1$, then p **is causally sufficient** for $X = 1$ (similarly for $X = 0$).

The Rachel Maddow example, again

MADDOW: You worked for [the health insurance company] CIGNA for 15 years, you left last year. What **caused** you to change your mind about what you were doing and to leave?

POTTER: Well, two things. One, it was kind of gradually. One instance or in one regard because I was becoming increasingly skeptical of the kinds of insurance policies that the big insurance companies are promoting and marketing these days. [. . .]

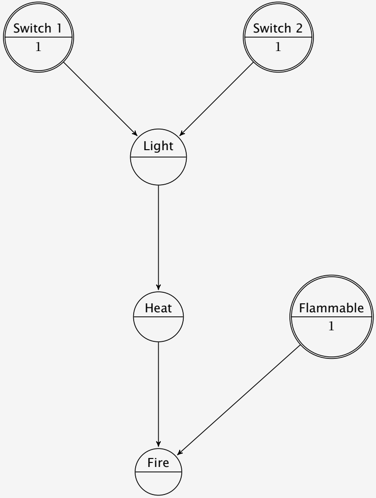
The other thing that really **made** me make this final decision to leave the industry occurred when I was visiting family in Tennessee a couple of summers ago, and [narrates the experience of happening on a 'healthcare expedition' where uninsured patients were treated by volunteer doctors in animal stalls at a fairground].

Why this is the wilderness

- There are some things that are very weird about this set-up.
- *e.g.*, it may be too liberal to allow background variables to change randomly.
- More generally, there is an unintended interdependence of the structure of the causal graph and time.
  - This can lead to 'race conditions'.

Extended Lifschitz example, again

Suppose, at m:



- f_{Light} as before.
- $f_{\text{Heat}} =$

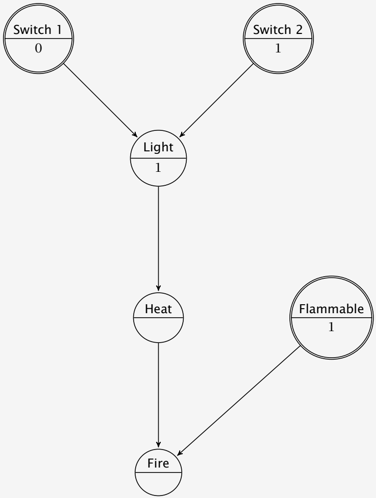
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example, again

Suppose, at $m+1$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

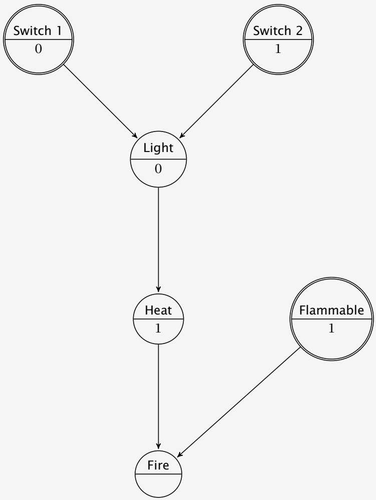
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example, again

Suppose, at $m + 2$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

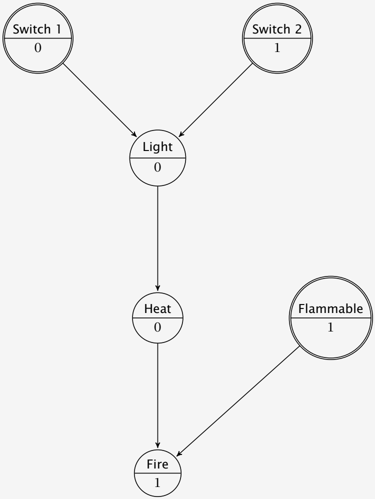
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example, again

Suppose, at $m+3$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

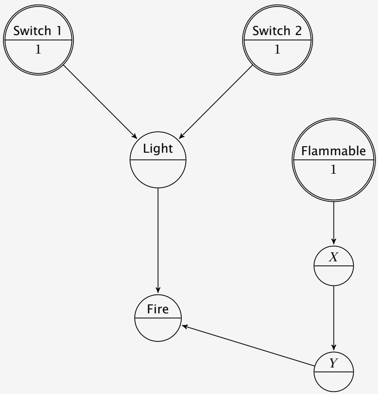
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example, once again

Suppose, at m:



- f_{Light} as before.
- f_{Heat} =

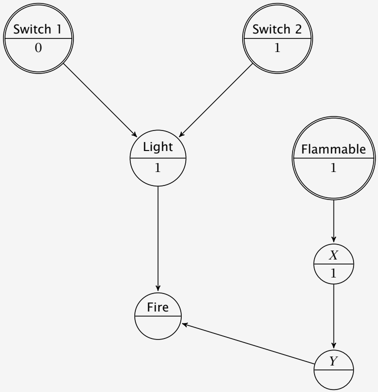
Light	Heat
0	0
1	1

- f_{Fire} =

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example, once again

Suppose, at $m + 1$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

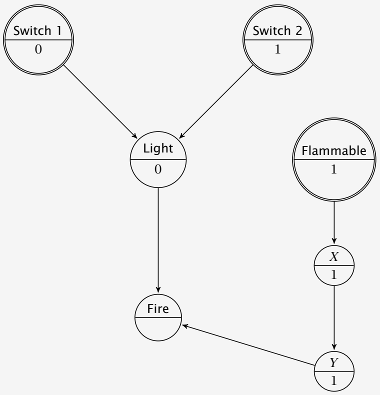
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example, once again

Suppose, at $m + 2$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

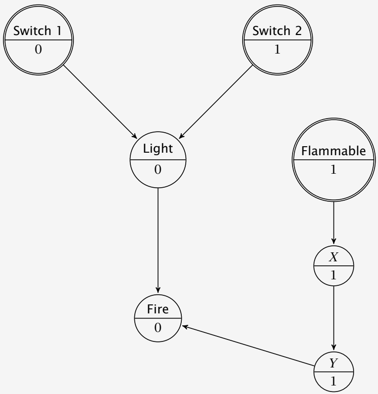
Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Extended Lifschitz example, once again

Suppose, at $m + 3$:



- f_{Light} as before.
- $f_{\text{Heat}} =$

Light	Heat
0	0
1	1

- $f_{\text{Fire}} =$

Heat	Flammable	Fire
0	0	0
0	1	0
1	0	1
1	1	0

Conclusion

- What I have: An analytical intuition that is promising.
  - Periphrastic causative verbs differ with respect to whether they assert causal necessity (*cause*?) or sufficiency.
  - Sufficiency causatives can appear to assert necessity by a pragmatic inference (causal perfection).
  - The coercive implication of *make* should follow from the sufficiency entailment.
- What I (still) need: A way to spell out causal sufficiency that gives us the coercive implication (when appropriate).
- Open questions:
  - How does *cause* work? Is it correct that it entails necessity? Does it also entail sufficiency? Or is this 'reverse perfection'?
  - What about other causative(-like) predicates? *let*? *allow*? *permit*? *lassen*?

Baglini, R. and Francez, I.: 2015, The implications of managing, *Journal of Semantics* **33**(3), 541–560.

Franke, M.: 2009, *Signal to Act: Game Theory and Pragmatics*, PhD thesis, Universiteit van Amsterdam, Amsterdam, The Netherlands.

Geis, M. L. and Zwicky, A. M.: 1971, On invited inferences, *Linguistic Inquiry* **2**, 561–566.

Horn, L. R.: 2000, From if to iff: Conditional perfection as pragmatic strengthening, *Journal of Pragmatics* **32**, 289–326.

Lewis, D.: 1973, Causation, *Journal of Philosophy* **70**, 556–567.

Nadathur, P.: 2016, Causal necessity and sufficiency in implicativity, *Proceedings of Semantics and Linguistic Theory (SALT) 26*, pp. 1002–1021.

Schulz, K.: 2007, *Minimal Models in Semantics and Pragmatics. Free Choice, Exhaustivity, and Conditionals*, PhD thesis, Universiteit van Amsterdam.

Schulz, K.: 2011, "if you'd wiggled a, then b would've changed": Causality and counterfactual conditionals, *Synthese* **179**, 239–251.

van der Auwera, J.: 1997, Conditional perfection, *in* A. Athanasiadou and R. Dirven (eds), *On Conditionals Again*, John Benjamins, Amsterdam, pp. 169–190.

von Fintel, K.: 2001, Conditional strengthening: A case study in implicature. Manuscript, Massachusetts Institute of Technology.