A Characterization of (Weakly Centered) Lewisian Causal Models

Jiji Zhang

Lingnan University, Hong Kong

Asian Logic Conference, June 21, 2019

Outline

- Interventionist conditionals and causal model semantics
- Stalnaker-Lewis ordering semantics
- Characterizations of 3 classes of causal models
  - Stalnakerian
  - Lewisian
  - Weakly-centered Lewisian

Interventionist conditionals

- In the interventionist approach to causation and causal explanation (Woodward 2003), the relevant counterfactual conditionals are interpreted as stating consequences of hypothetical interventions.

- The best known formal semantics for such interventionist conditionals is due to Pearl (2000) and Halpern (2000), which is formulated in terms of Functional Causal Models (a.k.a Structural Equation Models).

Functional causal model

- Following Halpern (2000), a signature is a tuple $<\mathbf{U}, \mathbf{V}, R>$, where **U** and **V** are finite sets of variables, and R associates with each variable $X \in \mathbf{U} \cup \mathbf{V}$ a finite set of values $R(X)$.

- A (functional) causal model over a signature $S = <\mathbf{U}, \mathbf{V}, R>$ is a tuple $<S, F>$, where F is a collection of functions, one for each $X \in \mathbf{V}$, $f_X \colon \times_{Y \in \mathbf{U} \cup \mathbf{V} \setminus \{X\}} R(Y) \to R(X)$.

- Notice that there are no functions for the variables in **U**, which are called *exogenous*. (Variables in **V** are called *endogenous*.) A value assignment **u** of **U** is called a *context*.

Causal graph

- A causal model <**U**, **V**, *R*, *F*> is associated with a directed graph: where each $X \in \mathbf{U} \cup \mathbf{V}$ is represented as a node, and there is an arrow from Y to X iff Y is a non-redundant argument in f_X.

- The model is called <u>recursive</u> iff the associated graph is acyclic.

Example

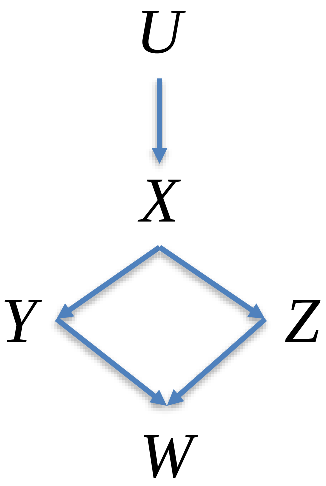
U={U}; **V**={X, Y, Z, W}; all Boolean variables.

$$X = U$$

$$Y = X$$

$$Z = X$$

$$W = max(Y, Z)$$



Intervention

Given $M = \langle \mathbf{U}, \mathbf{V}, R, F \rangle$, $\mathbf{X} \subseteq \mathbf{V}$, and a value assignment $\mathbf{x}$ of $\mathbf{X}$,

The *submodel of M with respect to* $\mathbf{X}=\mathbf{x}$, denoted by $M[\mathbf{X}=\mathbf{x}]$, is $\langle \mathbf{U}, \mathbf{V}, R, F_{\mathbf{X}=\mathbf{x}} \rangle$, where $F_{\mathbf{X}=\mathbf{x}}$ differs from F only for variables in $\mathbf{X}$: for every $X \in \mathbf{X}$, the function in $F_{\mathbf{X}=\mathbf{x}}$ for X is the constant function $X=x$, where x is the component value of X in $\mathbf{x}$.

- $M[\mathbf{X}=\mathbf{x}]$ is meant to model an external intervention that enforces $\mathbf{X}=\mathbf{x}$.

- If $\mathbf{X}=\varnothing$ (and so $\mathbf{x}$=nil), then $M[\mathbf{X}=\mathbf{x}] = M$.

Example

$\mathbf{U}=\{U\}$; $\mathbf{V}=\{X, Y, Z, W\}$; all Boolean variables.

$$X = U$$
$$Y = X$$
$$Z = X$$
$$W = max(Y, Z)$$

An intervention that fixes the value of Y to 1 is then represented by:

$$X = U$$
$$\sout{Y = X} \quad Y = 1$$
$$Z = X$$
$$W = max(Y, Z)$$

Syntax

- Following Halpern (2000, 2013), we consider a simple language over a signature $<\mathbf{U}, \mathbf{V}, R>$:

$$X{=}x \mid {\sim}\alpha \mid \alpha{\wedge}\beta \mid \alpha{\supset}\beta \mid \mathbf{X}{=}\mathbf{x} \;\Box{\rightarrow}\; \eta,$$

where

(1) $X \in \mathbf{V}$ and $x \in R(X)$;

(2) $\mathbf{X}{=}\mathbf{x}$ stands for a wff of the form $(X_1{=}x_1) \wedge \ldots \wedge (X_n{=}x_n)$, s.t. all X_i's are distinct, and η is a wff that does not contain '$\Box{\rightarrow}$'.

- No iterated conditionals and no conditionals with disjunctive antecedents (Briggs, 2012).

Semantics

- Let $M = \langle \mathbf{U}, \mathbf{V}, R, F \rangle$ and $\mathbf{u}$ be a context. A *solution* to M relative to $\mathbf{u}$ is a value assignment $\mathbf{v}$ of $\mathbf{V}$ such that $\mathbf{u}$ and $\mathbf{v}$ are consistent with all the functions in F.
- Assume M is <u>solutionful</u> in the sense that it has at least one solution relative to every context.
- Given M and any solution $(\mathbf{u}, \mathbf{v})$ relative to a context $\mathbf{u}$:

$\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models X{=}x$ iff $\mathbf{v}$ assigns value x to X.

$\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models {\sim}\alpha$ iff $\langle M, (\mathbf{u}, \mathbf{v}) \rangle \not\models \alpha$.

$\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models \alpha \wedge \beta$ iff $\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models \alpha$ and $\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models \beta$.

$\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models \alpha \supset \beta$ iff $\langle M, (\mathbf{u}, \mathbf{v}) \rangle \not\models \alpha$ or $\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models \beta$.

$\langle M, (\mathbf{u}, \mathbf{v}) \rangle \models \mathbf{X}{=}\mathbf{x} \mathbin{\Box\!\!\rightarrow} \eta$ iff for every solution $(\mathbf{u}, \mathbf{v}')$ to $M[\mathbf{X}{=}\mathbf{x}]$ relative to $\mathbf{u}$, $\langle M[\mathbf{X}{=}\mathbf{x}], (\mathbf{u}, \mathbf{v}') \rangle \models \eta$.

Pearl's constraint on causal models

- In Pearl (2000)'s definition of causal models, a "unique-solution" condition is imposed: every submodel has a unique solution relative to every context.

- Denote the set of Pearlian models by $\boldsymbol{M}_{\mathrm{uniq}}$, and the set of recursive models by $\boldsymbol{M}_{\mathrm{rec}}$. It is easy to show that $\boldsymbol{M}_{\mathrm{rec}} \subset \boldsymbol{M}_{\mathrm{uniq}}$.

Question

How is the causal-model semantics related to the Stalnaker-Lewis ordering semantics (for the simple language)?

Ordering semantics

- An <u>ordering model</u> over a signature S = <**U**, **V**, R> is a tuple <S, Ω, O, π>, where
  - Ω is a (finite) set of worlds;
  - π assigns a value $x \in R(X)$ to X at w, for every $X \in \mathbf{U} \cup \mathbf{V}$ and $w \in \Omega$;
  - O associates with each $w \in \Omega$ a subset of worlds, $\Omega_w \subseteq \Omega$ such that $w \in \Omega_w$, and a weak order $\leq_w$ over Ω_w.

- Given an ordering model $\mathcal{M}$ = <S, Ω, O, π> and a world $w \in \Omega$,

$<\mathcal{M}, w> \models X=x$ iff $\pi(X, w) = x$.

…

$<\mathcal{M}, w> \models \mathbf{X}=\mathbf{x} \;\Box\!\!\rightarrow \eta$ iff for every $w' \in C_{\mathcal{M}}(w, \mathbf{X}=\mathbf{x})$, $<\mathcal{M}, w'> \models \eta$,

where $C_{\mathcal{M}}(w, \mathbf{X}=\mathbf{x}) = \{v \mid v \in \Omega_w, \pi(\mathbf{X}, v) = \mathbf{x}$, and $v \leq_w v'$ for every $v' \in \Omega_w, \pi(\mathbf{X}, v') = \mathbf{x}\}$.

Constraints

- An ordering model $<S, \Omega, O, \pi>$ is called <u>weakly-centered Lewisian</u> if for every $w \in \Omega$ and every $v \in \Omega_w$, $w \leq_w v$.

- An ordering model $<S, \Omega, O, \pi>$ is called <u>Lewisian</u> if for every $w \in \Omega$ and every $v \in \Omega_w$ such that $v \neq w$, $w <_w v$ (i.e, $w \leq_w v$ but not $v \leq_w w$).

- An ordering model $<S, \Omega, O, \pi>$ is called <u>Stalnakerian</u> if it is Lewisian and for every $w \in \Omega$, $\leq_w$ is a linear order.

- Call a causal model *Lewisian (weakly-centered Lewisian*, *Stalnakerian)* if there is a Lewisian (weakly-centered Lewisian, Stalnakerian) ordering model that validates the exact same sentences (in the present language) as the causal model does.

Halpern's result

- Halpern (2013) showed that every recursive causal model is Lewisian, but not every Pearlian causal model is.

- Halpern suggested the following moral: "My own feeling is that these arguments show that models in $\boldsymbol{M}_{\text{uniq}}$-$\boldsymbol{M}_{\text{rec}}$ are actually not good models for causality. … I am not aware of any interesting real-world situation that is captured by a model in $\boldsymbol{M}_{\text{uniq}}$-$\boldsymbol{M}_{\text{rec}}$."

- An (oversimplified) Cobweb model:

$$P = \text{a} - \text{b}Q + U_1$$

$$Q = \text{c} + \text{d}P + U_2$$

Some conditions on causal models

- A causal model M is said to be Solutionful if M has a solution relative to every context.

- A causal model M is said to be Solution-Determinate if M has at most one solution relative to every context.

- A causal model M is said to be Solution-Conservative if for every $X \in \mathbf{V}$, every $x \in R(X)$, and every context $\mathbf{u}$, if M has a solution relative to $\mathbf{u}$ that assigns value x to X, then every solution to $M[X=x]$ relative to $\mathbf{u}$ is also a solution to M relative to $\mathbf{u}$ (Zhang 2013).

One more

- A causal model M is said to be Solution-Transitive in Cycles if for every $\mathbf{X_1}, \ldots, \mathbf{X_k} \subseteq \mathbf{V}$, every value $\mathbf{x_i}$ of $\mathbf{X_i}$, and every context $\mathbf{u}$, if for $1 \leq i \leq k$, $M[\mathbf{X_i}=\mathbf{x_i}]$ has a solution relative to $\mathbf{u}$ that is consistent with $\mathbf{X_{i+1}} = \mathbf{x_{i+1}}$ (where $\mathbf{X_{k+1}}$ denotes $\mathbf{X_1}$ and $\mathbf{x_{k+1}}$ denotes $\mathbf{x_1}$), then $M[\mathbf{X_1}=\mathbf{x_1}]$ has a solution relative to $\mathbf{u}$ that is consistent with $\mathbf{X_k}=\mathbf{x_k}$.

Characteristic axiom scheme:

$$(\mathbf{X_1}=\mathbf{x_1} \Diamond\!\!\rightarrow \mathbf{X_2}=\mathbf{x_2}) \wedge \ldots \wedge (\mathbf{X_k}=\mathbf{x_k} \Diamond\!\!\rightarrow \mathbf{X_1}=\mathbf{x_1})$$
$$\supset (\mathbf{X_1}=\mathbf{x_1} \Diamond\!\!\rightarrow \mathbf{X_k}=\mathbf{x_k})$$

where $\alpha \Diamond\!\!\rightarrow \beta =_{df} \sim(\alpha \Box\!\!\rightarrow \sim\beta)$.

Characterizations

Theorem: A causal model *M* is Stalnakerian iff *M* is Solution-Transitive in Cycles and is Pearlian (i.e., every submodel of *M* is Solutionful and Solution-determinate.)

Theorem: A causal model *M* is Lewisian iff *M* is Solution-Transitive in Cycles and Solution-determinate, and every submodel of *M* is Solutionful and Solution-Conservative.

Theorem: A causal model *M* is weakly-centered Lewisian iff *M* is Solution-Transitive in Cycles, and every submodel of *M* is Solutionful and Solution-Conservative.

Independent justifications?

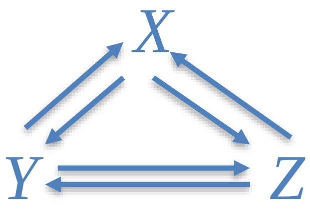
- Solution-Conservativeness: An intervention that fixes a variable to its value in a “natural equilibrium” (an equilibrium that might result without any intervention) should not lead to any new equilibrium.

- Solution-Transitivity in Cycles: There should be no cycle of counterfactual dependence (Zhang et al., 2013).

An example of a "bad" model

- Consider a non-recursive causal model with 3 binary variables (all of which are endogenous):

$$X = max(1\text{-}Y, Z); \quad Y = max(1\text{-}Z, X); \quad Z = max(1\text{-}X, Y)$$



- This model is not Solution-Transitive in Cycles and so not Lewisian. It satisfies:

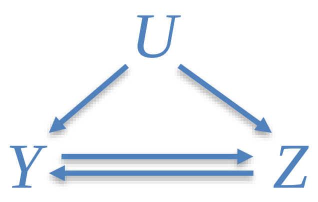
$$(X=1 \;\Box\!\!\rightarrow Y=1) \wedge (X=0 \;\Box\!\!\rightarrow Y=0)$$
$$(Y=1 \;\Box\!\!\rightarrow Z=1) \wedge (Y=0 \;\Box\!\!\rightarrow Z=0)$$
$$(Z=1 \;\Box\!\!\rightarrow X=1) \wedge (Z=0 \;\Box\!\!\rightarrow X=0)$$

That is, a cycle of counterfactual dependence.

An example of a not-so-bad model

- Consider a non-recursive causal model with 3 binary variables (two endogenous, one exogenous):

$$Y = max(U, Z),\ \ Z = max(1\text{-}U, Y)$$



- Although this model is non-recursive, it satisfies Solution-Transitivity in Cycles and hence does not feature any cycle of counterfactual dependence.

- So a Lewisian constraint on causal models is that there is no cycle at the level of values, though there might be cycles at the level of variables. (Type vs token? Determinable vs determinate?)

Conclusions

- The class of recursive causal models is far from the full class of Lewisian causal models.

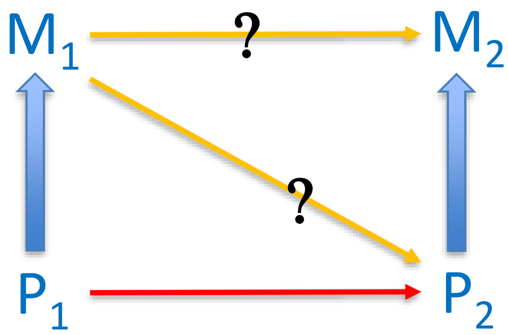
- Some Lewisian constraints on causal models seem to be interesting on independent grounds.

- The permission of (non-recursive) models with multiple solutions naturally relaxes Centering to Weak Centering.

Some references

- Briggs, R. (2012). "Interventionist counterfactuals". Philosophical Studies, 160:139–166.
- Halpern, J. Y. (2000). Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research* 12:317-37.
- Halpern, J. Y. (2013). From causal models to counterfactual structures. *The Review of Symbolic Logic* 6(2):305-22.
- Pearl, J. (2000). *Causality*: *Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.
- Zhang, J. (2013). A Lewisian logic of causal counterfactuals. *Minds and Machines*, 23(1): 77-93.
- Zhang, J., W. Lam, and R. De Clercq. (2013). A peculiarity in Pearl's logic of interventionist counterfactuals. *Journal of Philosophical Logic*, 42(5):783-94.

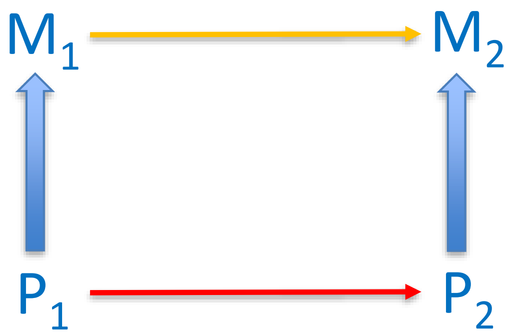
Exclusion and Weak Centering



- Kim (1998): M_1 cannot cause M_2 (or P_2), for P_1 is already causally sufficient.

- List and Menzies (2009): M_1 can be a cause of M_2, if one adopts a counterfactual dependence conception of causation and weaken Lewis's constraints to require only Weak Centering instead of Centering.

Exclusion and Weak Centering



- Zhong (2014): Causal autonomy (as depicted in the diagram) is possible if one adopts an interventionist conception of causation.

- Sze (2015): Zhong's argument implicitly assumes an understanding of interventionist counterfactuals that abandons Centering in favour of Weak Centering.

Acknowledgements

This research was supported by Hong Kong Research Grants Council, General Research Fund LU13602818.