A REPRESENTATION THEOREM FOR CAUSAL DECISION MAKING

Evan Piermont
Royal Holloway, University of London

Theory Workshop

June 2024

The ⟶ of time

- Time appears to advance in a single direction, from earlier to later
- Causality is the influence of earlier events on later ones
- What exactly constitutes a *causal relation* is philosophically sticky
    - Taking the structure causality as given, identifying causal relations is still not straightforward

This paper

We represent causality via *structural equations*, and consider an agent's preference over *interventions*:

- Representation Theorem
  - How an agent's subjective causal model influence her decision making
- Identification Theorem
  - When can this model be recovered from observation

Causation and Counterfactuals

- Modern theories define causation through counterfactuals.
- Simplest form is ‘*but for*’ causality: α causes β
    - when α occurs so does β occur
    - had α not occurred, β would not occur
- There are many subtitles here
- Requires evaluating worlds that do not exist

Causation in Economics

Reduced Form

- At the population level
- Understood via conditional dependence
- I.e., Smoking causes cancer

Structural Form

- At the individual level
- Understood via equations between variables
- I.e., agent’s education level caused her earnings

Structural causality + uncertainty/hidden variables = reduced from causality

We take a structural approach a la Pearl [2000]:

- Equations relate the values of variables
- These equations directly encode causal mechanisms
- Provide a succinct way of contemplating counterfactuals

Causal Models

A *causal model* **M** consists of:

- $\mathcal{U}$ and $\mathcal{V}$ denote exogenous and endogenous variables, resp.
- $\mathcal{R}(Z) \subset \mathbb{R}$ is the range of $Z \in \mathcal{U} \cup \mathcal{V}$
- $\mathcal{F} = \{F_X\}_{X \in \mathcal{V}}$ is a set of **structural equations** where

$$F_X : \prod_{Y \in \mathcal{U} \cup (\mathcal{V} - \{X\})} \mathcal{R}(Y) \to \mathcal{R}(X).$$

Causal Models

- Call **M** *recursive* if
  - exists a partial order on $\mathcal{V}$
  - F_X is independent of the variables succeeding X

Causal Models

- A *context* is a vector $\vec{u}$ of values for all the exogenous variables $\mathcal{U}$.
  - Let $\texttt{ctx} = \prod_{U \in \mathcal{U}} \mathcal{R}(U)$ collect all contexts
- A *resolution* is a vector $\vec{a}$ of values for all variables $\mathcal{U} \cup \mathcal{V}$.
  - Let $\texttt{res} = \prod_{Y \in \mathcal{U} \cup \mathcal{V}} \mathcal{R}(Y)$ collect all resolutions
- Given a recursive model, each context $\vec{u}$ uniquely determines a resolution $\vec{a}$.

Example

The US Federal Reserve is contemplating the economy.

The relevant variables are: the growth rate (grow), the prior interest rate (prior), the current interest rate (rate), inflation (inf), employment rate (emp):

$$\mathcal{U} = \begin{cases} U_{grow} \\ U_{prior} \end{cases} \qquad \mathcal{V} = \begin{cases} Y_{rate} \\ X_{emp} \\ X_{inf} \end{cases}$$

Example

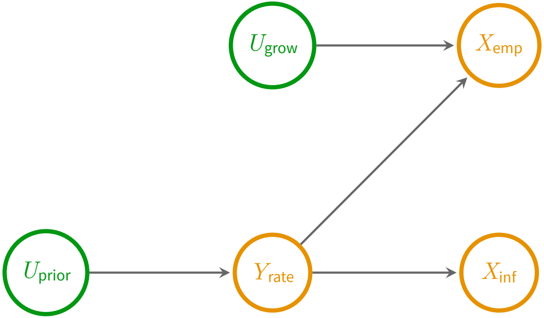
Assume for simplicity that all variables take values in in $\{0, 1\}$. The causal equations are

$$X_{inf} = 1 - Y_{rate} \qquad (F_{X_{inf}})$$

$$X_{emp} = 1 - (Y_{rate} \times (1 - U_{grow})) \qquad (F_{X_{inf}})$$

$$Y_{rate} = U_{prior} \qquad (F_{Y_{rate}})$$

Example



Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$
\begin{aligned}
Y_{rate} &= U_{prior} \\
X_{inf} &= 1 - Y_{rate} \\
X_{emp} &= 1 - (Y_{rate} \times (1 - U_{grow}))
\end{aligned}
$$

Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$\begin{aligned} Y_{rate} &= 0 \\ X_{inf} &= 1 - Y_{rate} \\ X_{emp} &= 1 - (Y_{rate} \times (1 - 0)) \end{aligned}$$

Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$\begin{aligned} Y_{rate} &= \ 0 \\ X_{inf} &= \ 1 - 0 \\ X_{emp} &= \ 1 - (0 \times (1 - 0)) \end{aligned}$$

Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$\begin{aligned} Y_{rate} &= \ 0 \\ X_{inf} &= \ 1 \\ X_{emp} &= \ 0 \end{aligned}$$

Example

If instead both were high: $\vec{u} = \begin{cases} U_{grow} = 1 \\ U_{prior} = 1 \end{cases}$

$$\begin{aligned} Y_{rate} &= U_{prior} \\ X_{inf} &= 1 - Y_{rate} \\ X_{emp} &= 1 - (Y_{rate} \times (1 - U_{grow})) \end{aligned}$$

Example

If instead both were high: $\vec{u} = \begin{cases} U_{grow} = 1 \\ U_{prior} = 1 \end{cases}$

$$\begin{aligned} Y_{rate} &= 1 \\ X_{inf} &= 1 - Y_{rate} \\ X_{emp} &= 1 - (Y_{rate} \times (1 - 1)) \end{aligned}$$

Example

If instead both were high: $\vec{u} = \begin{cases} U_{grow} = 1 \\ U_{prior} = 1 \end{cases}$

$$
\begin{aligned}
Y_{rate} &= \ 1 \\
X_{inf} &= \ 1 - 1 \\
X_{emp} &= \ 1 - (1 \times (1 - 1))
\end{aligned}
$$

Example

If instead both were high: $\vec{u} = \begin{cases} U_{grow} = 1 \\ U_{prior} = 1 \end{cases}$

$$\begin{aligned} Y_{rate} &= 1 \\ X_{inf} &= 0 \\ X_{emp} &= 1 \end{aligned}$$

Interventions & Actions

A *intervention*

$$\mathbf{do}[\,Y_1 \leftarrow y_1, \ldots, Y_n \leftarrow y_n]$$

is a mediation that sets the values of $Y_1 \ldots Y_n \in \mathcal{V}$:

- $y_i \in \mathcal{R}(Y_i)$
- abbreviated as $\mathbf{do}[\,\vec{Y} \leftarrow \vec{y}]$
- interventions only on endogenous variables.

Interventions & Actions

The *intervention*

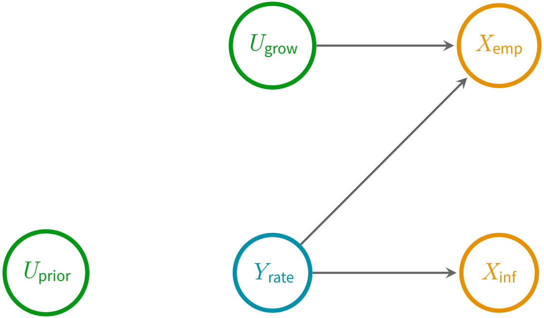
$$\mathbf{do}[\,Y_1 \leftarrow y_1, \ldots, Y_n \leftarrow y_n]$$

induces a *counterfactual model*, $\mathcal{F}_{\mathbf{do}[\,\vec{Y} \leftarrow \vec{y}]}$ where

F_{Y_i} is replaced by the constant function $F'_{Y_i} = y_i$

Example

The action $\mathbf{do}[Y_{rate} \leftarrow 1]$ sets the current rate to 1:



Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$\begin{aligned} Y_{rate} &= 1 \\ X_{inf} &= 1 - Y_{rate} \\ X_{emp} &= 1 - (Y_{rate} \times (1 - U_{grow})) \end{aligned}$$

Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$\begin{aligned} Y_{rate} &= 1 \\ X_{inf} &= 1 - Y_{rate} \\ X_{emp} &= 1 - (Y_{rate} \times (1 - 0)) \end{aligned}$$

Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$\begin{aligned} Y_{rate} &= \; 1 \\ X_{inf} &= \; 1 - 1 \\ X_{emp} &= \; 1 - (1 \times (1 - 0)) \end{aligned}$$

Example

Given the context: $\vec{u} = \begin{cases} U_{grow} = 0 \\ U_{prior} = 0 \end{cases}$

$$\begin{aligned} Y_{rate} &= 1 \\ X_{inf} &= 0 \\ X_{emp} &= 0 \end{aligned}$$

Example

$$U_{grow} = 0$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 1]} \left\{ \begin{aligned} Y_{rate} &= 1 \\ X_{inf} &= 0 \\ X_{emp} &= 0 \end{aligned} \right.$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 0]} \left\{ \begin{aligned} Y_{rate} &= 0 \\ X_{inf} &= 1 \\ X_{emp} &= 1 \end{aligned} \right.$$

$$U_{grow} = 1$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 1]} \left\{ \begin{aligned} Y_{rate} &= 1 \\ X_{inf} &= 0 \\ X_{emp} &= 1 \end{aligned} \right.$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 0]} \left\{ \begin{aligned} Y_{rate} &= 0 \\ X_{inf} &= 1 \\ X_{emp} &= 1 \end{aligned} \right.$$

Actions

An *action* is of the form:

$$\textbf{if } \phi \textbf{ then } A \textbf{ else } B$$

- ϕ is a true/false valued question about the variable values
    - such as “the value of X is positive”, etc
- A and B are actions
- These is constructed recursively starting with interventions
- **if** ϕ **then** A shorthand for when $B = \emptyset$

Actions

Given a (recursive) model **M** and action A, let

$$\beta_A^{\mathbf{M}} : \mathtt{ctx} \to \mathtt{res}$$

transform contexts into resolutions in the obvious way:

- Each context determines which 'clause' of A will be in force, hence an intervention
- This intervention determines a (recursive) counterfactual model
- Along with context, this determines the resolution

Preference

The observable of the model is an agent's preference relation $\succsim$ over actions

Representation

The agent's preferences are parameterized by

- $\mathbf{M}$ — a recursive model capturing causal relationships
- $\mathbf{u} : \mathtt{res} \to \mathbb{R}$ — value of a resolution of all uncertainty
- $\mathbf{p} \in \Delta(\mathtt{ctx})$ — belief capturing uncertainty about the values of exogenous (hence endogenous) variables

Representation

Subjective Causal Utility

$(\mathbf{M}, \mathbf{p}, \mathbf{u})$ is a **subjective causal utility representation** of $\succsim$:

$$A \succsim B$$

if and only if

$$\sum_{\vec{u} \in \mathtt{ctx}} \mathbf{u}(\beta_A^{\mathbf{M}}(\vec{u}))\mathbf{p}(\vec{u}) \geq \sum_{\vec{u} \in \mathtt{ctx}} \mathbf{u}(\beta_B^{\mathbf{M}}(\vec{u}))\mathbf{p}(\vec{u}).$$

Example

The utility of the Federal Reserve is determined by the inflation rate and employment level, and is given by

$$\mathbf{u}(\vec{a}) = 2X_{emp} - X_{inf}.$$

Example

$$U_{grow} = 0$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 1]} \begin{cases} Y_{rate} & = 1 \\ X_{inf} & = 0 \\ X_{emp} & = 0 \end{cases}$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 0]} \begin{cases} Y_{rate} & = 0 \\ X_{inf} & = 1 \\ X_{emp} & = 1 \end{cases}$$

$$U_{grow} = 1$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 1]} \begin{cases} Y_{rate} & = 1 \\ X_{inf} & = 0 \\ X_{emp} & = 1 \end{cases}$$

$$\beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 0]} \begin{cases} Y_{rate} & = 0 \\ X_{inf} & = 1 \\ X_{emp} & = 1 \end{cases}$$

Example

$U_{grow} = 0$

$$\mathbf{u} \circ \beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 1]} 2 \times 0 - 0 = 0$$

$$\mathbf{u} \circ \beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 0]} 2 \times 1 - 1 = 1$$

$$\mathbf{do}[Y_{rate} \leftarrow 0] \succsim \mathbf{do}[Y_{rate} \leftarrow 1]$$

$U_{grow} = 1$

$$\mathbf{u} \circ \beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 1]} 2 \times 1 - 0 = 2$$

$$\mathbf{u} \circ \beta \xrightarrow{\mathbf{do}[Y_{rate} \leftarrow 0]} 2 \times 1 - 1 = 1$$

$$\mathbf{do}[Y_{rate} \leftarrow 1] \succsim \mathbf{do}[Y_{rate} \leftarrow 0]$$

Example

- Preference between setting interest rate at 1 or 0 depends on belief about U_{grow}.
- The (conditional) action

$$\textbf{if}\ (U_{grow} = 1)\ \textbf{then}\ \textbf{do}[Y_{rate} \leftarrow 1]\ \textbf{else}\ \textbf{do}[Y_{rate} \leftarrow 0]$$

dominates

Axioms

Ax 1: Cancellation

Let $A_1 \dots A_n$ and $B_1 \dots B_n$ be actions such that, for all $\vec{u} \in \texttt{ctx}$ and interventions $\mathbf{do}[Y \leftarrow y]$ we have

$$\#\{A_i \mid A_i \text{ induces } \mathbf{do}[Y \leftarrow y] \text{ given } \vec{u}\}$$
$$=$$
$$\#\{B_i \mid B_i \text{ induces } \mathbf{do}[Y \leftarrow y] \text{ given } \vec{u}\}$$

then $A_i \succsim B_i$ for all $i < n$ implies $B_n \succsim A_n$.

- Adapted from Blume, Easley, Halpern (2021)
- Provides an (abstract) additive structure

`Ax 2: Model Uniqueness`

For each $\vec{u} \in \texttt{ctx}$, there is at most one $\vec{a} \in \texttt{res}$ such that $\vec{a}|_{\mathcal{U}} = \vec{u}$ and $\vec{a}$ is non-null.

- Non-null: $(\textbf{if } \vec{a} \textbf{ then } A) \succ (\textbf{if } \vec{a} \textbf{ then } B)$ for some A, B.
- The only uncertainty regards the context

For each $\vec{a} \in \texttt{res}$, write

$$\mathbf{do}[\vec{Y} \leftarrow \vec{y}] \rightsquigarrow_{\vec{a}} (X = x)$$

as shorthand for the indifference relation

$$\textbf{if } \vec{a} \textbf{ then } \mathbf{do}[\vec{Y} \leftarrow \vec{y}, X \leftarrow x] \sim \textbf{if } \vec{a} \textbf{ then } \mathbf{do}[\vec{Y} \leftarrow \vec{y}].$$

- If setting $\vec{Y}$ to $\vec{y}$ yields $X = x$, then the agent is indifferent from making such a further intervention on X.
- However, definition allows for indifference between distinct values of X

Ax 3: Definiteness

Fix non-null $\vec{a} \in \texttt{res}$, endogenous variables, $\vec{Y}$, and values $\vec{y} \in \mathcal{R}(\vec{Y})$. Then for variable X, there exists some $x \in \mathcal{R}(X)$ such that

$$\mathbf{do}[\vec{Y} \leftarrow \vec{y}] \rightsquigarrow_{\vec{a}} (X = x)$$

- There is some value of X which is consistent with any intervention
- May not be unique (i.e., indifference between resolutions)
- `Ax3`*: if the value x is unique

`Ax 4: Centeredness`

For $\vec{a} \in \texttt{res}$, vector of endogenous variables $\vec{Y}$, and endogenous variable $X \notin \vec{Y}$, we have

$$\mathbf{do}[\vec{Y} \leftarrow \vec{a}|_{\vec{Y}}] \rightsquigarrow_{\vec{a}} (X = \vec{a}|_X)$$

- Trivial interventions (setting variables to their current value) has no consequence

For X, $Y \in \mathcal{V}$, say that X *is unaffected by* Y if

$$\mathbf{do}[\vec{Z} \leftarrow \vec{z}] \rightsquigarrow_{\vec{a}} (X = x) \qquad \text{iff} \qquad \mathbf{do}[\vec{Z} \leftarrow \vec{z}, Y \leftarrow y] \rightsquigarrow_{\vec{a}} (X = x)$$

for all $\vec{a} \in \texttt{res}$, $\vec{Z}$ and values for the variables.

- X is unaffected by Y if there is no intervention on Y that changes the decision maker's perception of X
- If this relation does not hold, then X is *affected* by Y, written $Y \rightsquigarrow X$.

Ax 5: Recursivity

$\rightsquigarrow$ is acyclic

- There are no cycles of variable dependence

Theorem

$\succsim$ satisfies $\texttt{Ax1-5}$ if and only if there exists a subjective causal utility representation, $(\mathbf{M}, \mathbf{p}, \mathbf{u})$.

Moreover, if $\texttt{Ax3}^*$ holds, then $\mathbf{M}$ is unique.

Each axiom helps discipline how counterfactuals are constructed:

`Definiteness:`	There exists some counterfactual world
`Model Uniqueness:`	It is unique
`Centeredness:`	It is minimally different than the current world
`Recursivity:`	Closeness is consistent across contexts

These properties suffice to prove the existence of a structural model.

Thank You!