Abstracting Causal Models

Fabio Massimo Zennaro
fm.zennaro@gmail.com

University of Warwick
October 18th, 2021

Background

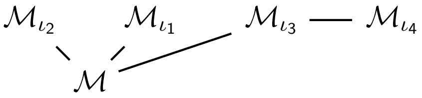
- We deal with **SCMs** defined as tuples $\langle \mathcal{X}, \mathcal{E}, \mathcal{F}, \mathcal{P} \rangle$ [1, 2, 3].

$$S \longrightarrow T \longrightarrow C$$

- We allow **perfect interventions** with $do(X_0 = x_0)$ operator [1, 2].

$$S \qquad 0 \longrightarrow C$$

- A SCM $\mathcal{M}$ serves as a **presentation** for a set of SCMs generated by interventions [4]:



Notation: we use $\mathcal{X}$ for a set of RV; X for a RV; $\mathcal{M}[X]$ for the domain of X

Problem definition

Suppose we are given two models $\mathcal{M}, \mathcal{M}'$ of the same phenomenon:

$$S \longrightarrow T \longrightarrow C \qquad\qquad S \longrightarrow C$$

What does it mean that *model* $\mathcal{M}'$ *is an* **abstraction** *of model* $\mathcal{M}$?

1. *Desideratum* of abstraction
2. *Formalization* of abstraction

Desideratum: commutativity

- **Abstraction-intervention commutativity:** given a model $\mathcal{M}$, the following two procedures lead to the same distribution $P_{\mathcal{M'}_{\iota'}}$:
  - Intervene on $\mathcal{M}$ and then map to the abstracted model;
  - Map $\mathcal{M}$ to the abstracted model and then intervene on it.

$$
\begin{tikzcd}
\mathcal{M} \arrow[r, dashed, "\alpha"] \arrow[d, "\iota"'] & \mathcal{M}' \arrow[d, "\iota'"] \\
\mathcal{M}_{\iota} \arrow[r, dashed, "\alpha"] & \mathcal{M}'_{\iota'}
\end{tikzcd}
$$

Formalization: statistical

- *Distributional*: α as a function mapping joint distributions [4]

$$\tau : \prod_{X \in \mathcal{X}} \mathcal{M}[X] \to \prod_{X' \in \mathcal{X}'} \mathcal{M}[X']$$

then we can assess:

$$\begin{array}{ccc} P_{\mathcal{M}} & \xrightarrow{\tau} & P_{\mathcal{M}'} \\ \Big\downarrow{\scriptstyle \iota} & & \Big\downarrow{\scriptstyle \iota'} \\ P_{\mathcal{M}_\iota} & \xrightarrow{\tau} & P_{\mathcal{M}'_{\iota'}} \end{array}$$

Formalization: statistical

- *Structural*: α as a collection of functions mapping variables [3]

$$
\begin{aligned}
R &\subseteq \mathcal{X} \\
a : &R \to \mathcal{X}' \\
\alpha_{X'} : &\mathcal{M}[a^{-1}(X')] \to \mathcal{M}[X']
\end{aligned}
$$

then we can assess:

$$
\begin{tikzcd}
 & \mathcal{M}[a^{-1}(X_1')] \arrow[r, "f"] \arrow[d, "\alpha_{X_1}"'] & \mathcal{M}[a^{-1}(X_2')] \arrow[d, "\alpha_{X_2}"] \\
\cdot \arrow[ur, "\iota"] \arrow[r, "\iota'"'] & \mathcal{M}[X_1'] \arrow[r, "f'"] & \mathcal{M}[X_2']
\end{tikzcd}
$$

Formalization: categorical

- *Structural*: diagrams such as

$$\begin{tikzcd}
\mathcal{M}[a^{-1}(X_1')] \arrow[r, "f"] \arrow[d, "\alpha_{X_1}"'] & \mathcal{M}[a^{-1}(X_2')] \arrow[d, "\alpha_{X_2}"] \\
\mathcal{M}[X_1'] \arrow[r, "f'"] & \mathcal{M}[X_2']
\end{tikzcd}$$

can live in the **FinStoch** category (*objects:* finite sets; *maps:* stochastic matrices).

FinStoch may be enriched in **Met** allowing for computation of *approximation error*.

Relevant research questions

- *Formalization:* how do different formalizations and perspectives on abstraction relate?
  - Distributional vs structural perspective
  - Preservation of causal structure
  - Furthering categorical formalization
- *Estimation:* how do we assess abstraction efficiently?
  - Choice of measure
  - Choice of interventions
  - Efficient algorithms

Further research questions

- *Causal representation learning:* can we integrate principle of abstraction in learning?
  - Guiding/explaining causal learning
- *Extensions:* what other aspects of abstractions between SCMs may be relevant?
  - Stochastic abstractions
  - Structure-preserving abstractions
  - Counterfactual consistency

High-level of interdisciplinarity (category theory, physics, graph theory).

Many interesting questions and promising directions!

Thanks!

Thank you for listening!

If interested in existing approaches, feel free to check tutorials at:
`https://github.com/FMZennaro/CategoricalCausalAbstraction`

References I

[1] Judea Pearl. *Causality*. Cambridge University Press, 2009.

[2] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: Foundations and learning algorithms*. MIT Press, 2017.

[3] Eigil Fjeldgren Rischel. The category theory of causal models. 2020.

[4] Paul K Rubenstein, Sebastian Weichwald, Stephan Bongers, Joris M Mooij, Dominik Janzing, Moritz Grosse-Wentrup, and Bernhard Schölkopf. Causal consistency of structural equation models. In *33rd Conference on Uncertainty in Artificial Intelligence (UAI 2017)*, pages 808–817. Curran Associates, Inc., 2017.