Learning Consistent Causal Abstractions (with Genetic Algorithms)

Fabio Massimo Zennaro

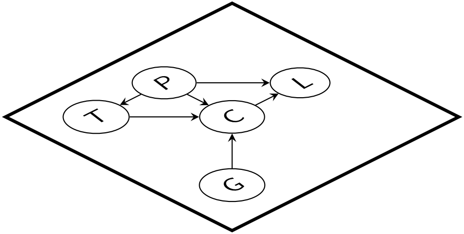
University of Bergen

September 10, 2024

1. Introduction

Causal Reasoning

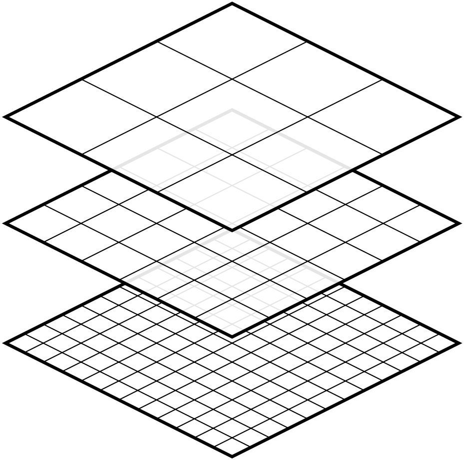
Causal reasoning is getting more relevant throughout ML/AI.



- It discriminates *correlations* from *causes*;
- It provides a *strong prior* for learning;
- It implies a *causality ladder* of reasoning;
- It offers *improved interpretability*.

Multilevel Reasoning

Multilevel/multiscale/multiresolution reasoning is common throughout the sciences.



- It allows for *multiple resolutions*;
- It aggregates *different observables*;
- It leads to *computational savings*;
- It allows shifting between *levels of abstraction*.

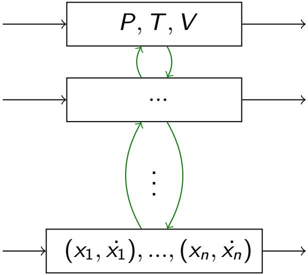
Levels of Abstraction

Systems may be represented at different **levels of abstraction** (LoA) [1].

Thermodynamics example:

Low-level / Base model:
Microscopic description $\mathbf{x}, \dot{\mathbf{x}}$.

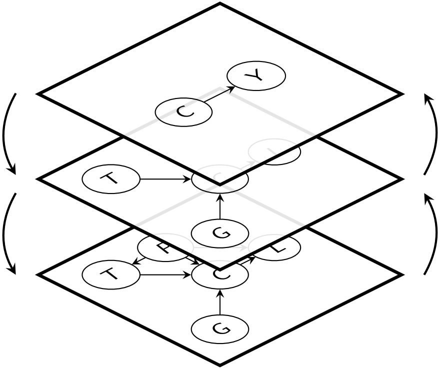
High-level / Abstracted model:
Macroscopic description P, T, V.



We want to be able to shift between LoAs **consistently**.

Causal Abstraction

Causal Abstraction joins *causal reasoning* and *multi-level reasoning*.



How do we *relate* causal models at different levels of abstraction?
How do we *learn* good abstractions?

2. Formalization

Formalization

Causality

C1. *Structural Causal Models*

C2. *Interventions*

Abstraction

A1. *Abstraction*

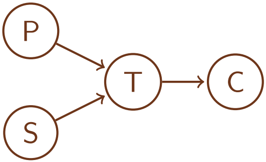
A2. α*-abstraction*

A3. *Interventional Consistency*

A4. *Abstraction Error*

C1. Structural Causal Models

We express a causal model as a **structural causal model** (SCM) $\mathcal{M} = \langle \mathcal{X}, \mathcal{U}, \mathcal{F}, \mathcal{P} \rangle$ [2, 3]:

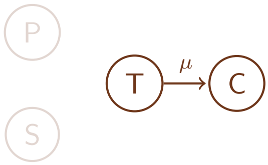


A **graphical model** with:

- a collection of *variables of interest*;
- a collection of *causal mechanisms*.

C2. Interventions

We can evaluate how **intervening** on a variable affects the system.



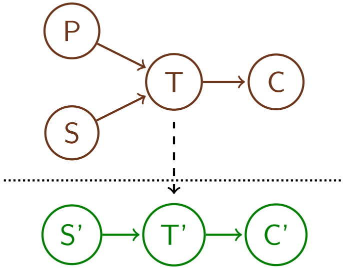
We can intervene (*do*):

- Set a variable (*cause*);
- Evaluate a distribution downstream (*effect*);

through a *mechanism* μ (*matrix*).

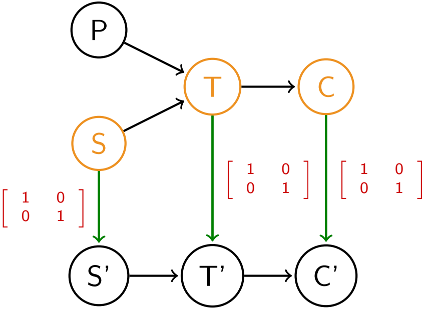
A1. Abstraction

We define an **abstraction** as a map between models.



A2. α-Abstraction

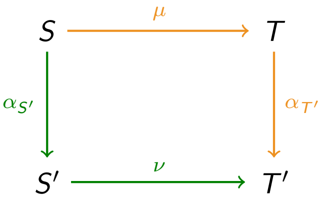
An **α-abstraction** $\langle R, a, \alpha_i \rangle$ [5, 4] is defined as:



- R: a set of *relevant variables*;
- a: a surjective function between *variables*;
- α_i: a collection of surjective functions between *outcomes* (*binary matrices*).

A3. Interventional consistency

We want an abstraction to guarantee *interventional consistency*.

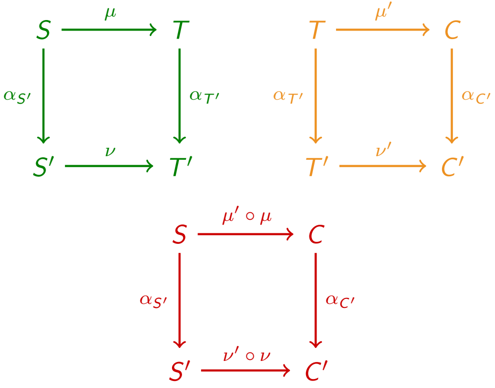


- Ideally, mechanisms and abstractions *commute*.
- Otherwise, we compute an abstraction error as the *worst-case discrepancy* over all possible interventions:

$$E_{\boldsymbol{\alpha}}(S', T') = \max D(\alpha_{T'} \cdot \mu, \nu \cdot \alpha_{S'})$$

A4. Abstraction Error

An abstraction implies multiple *causal mechanism diagrams*:



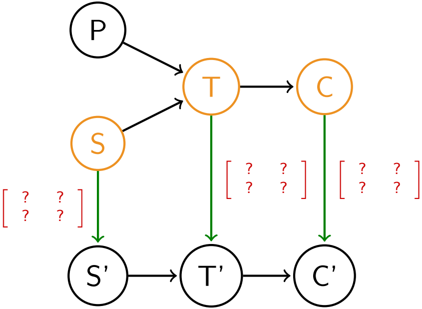
A **(global) abstraction error** [5] $e(\boldsymbol{\alpha})$ is the maximum abstraction error over all diagrams.

$$e(\boldsymbol{\alpha}) = \sup_{\mathbf{X}',\mathbf{Y}' \subseteq \mathcal{X}'} E_{\boldsymbol{\alpha}}(\mathbf{X}', \mathbf{Y}')$$

3. Problem Statement

Problem statement [6]

Given a partially defined *abstraction* $\boldsymbol{\alpha}$ in terms of $\langle R, a \rangle$ can we learn α_i?



Let's learn α_i as:

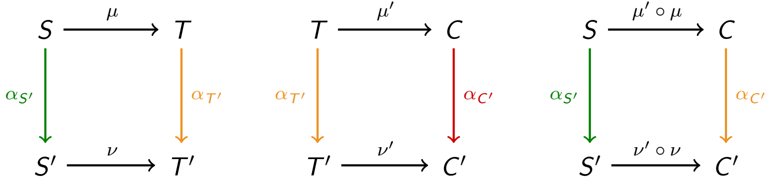
$$\min_{\boldsymbol{\alpha}} \; e(\boldsymbol{\alpha})$$

Challenges [6]

We need to learn multiple *maps/binary matrices*:

$$\alpha_{S'} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}, \alpha_{T'} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}, \alpha_{C'} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$$

while optimizing over multiple *diagrams*:



Several challenges:

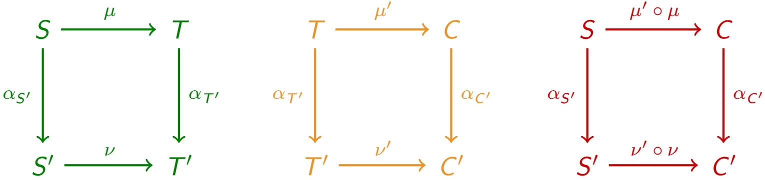
(i) *Multiple related problems*

(ii) *Combinatorial optimization*

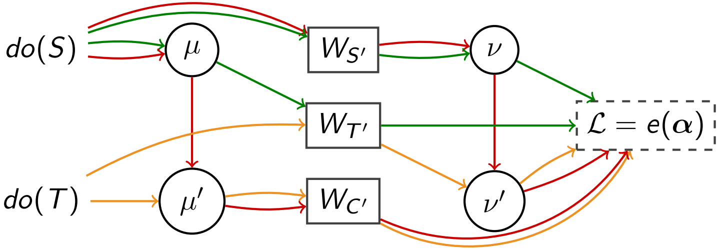
(iii) *Surjectivity constraints*

4. Solution Approaches

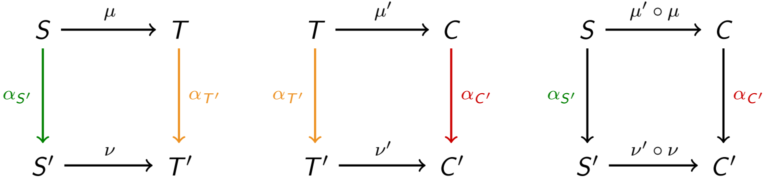
Solution by gradient descent [6]



Jointly solve all the problems via *relaxation* and *gradient descent*:



Solution by genetic algorithm



Encode the solutions in a *genotype* and define a *fitness* over all the problems, then solve by *genetic algorithms*:

$$f = -e(\boldsymbol{\alpha})$$

5. Conclusion

Conclusion

- *Causality* and *abstraction* both play an important role in modelling.
- *Causal abstraction* is relevant to:
    - *transportabiliy*;
    - *robustness*;
    - *interpretability*;
    - *causal representation learning*.
- It may be practically useful for *integrating data* and *reducing costs*.

Large space for conceptual and practical development of **causal abstraction frameworks**.

Thanks!

Thank you for listening!

More about causal abstraction:
https://github.com/FMZennaro/CausalAbstraction/

References I

[1] Luciano Floridi. The method of levels of abstraction. *Minds and machines*, 18(3):303–329, 2008.

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

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

[4] Eigil F Rischel and Sebastian Weichwald. Compositional abstraction error and a category of causal models. *arXiv preprint arXiv:2103.15758*, 2021.

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

[6] Fabio Massimo Zennaro, Máté Drávucz, Geanina Apachitei, W. Dhammika Widanage, and Theodoros Damoulas. Jointly learning consistent causal abstractions over multiple interventional distributions. In *2nd Conference on Causal Learning and Reasoning*, 2023.