Causal Models

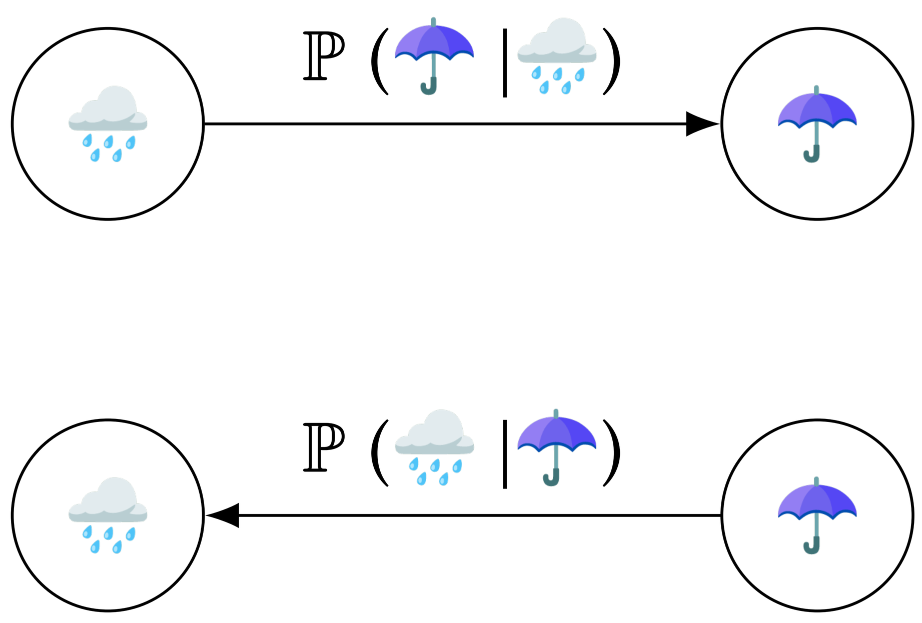
Learning, Representation, and Abstraction

Riccardo Massidda — PhD Candidate @ CS Department

Università di Pisa

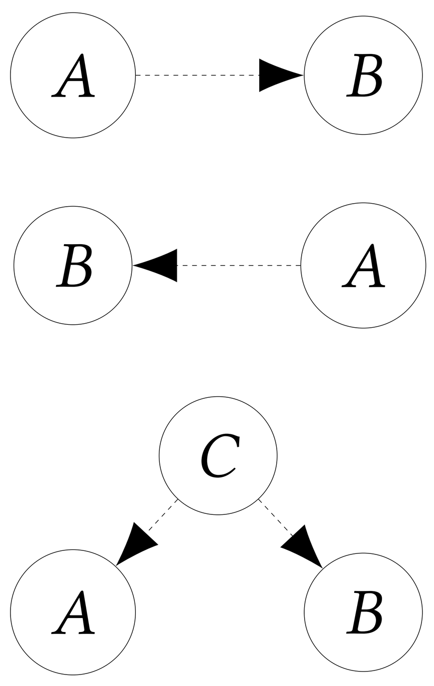
Causality

Causal information is fundamental to represent **manipulations** of a system.

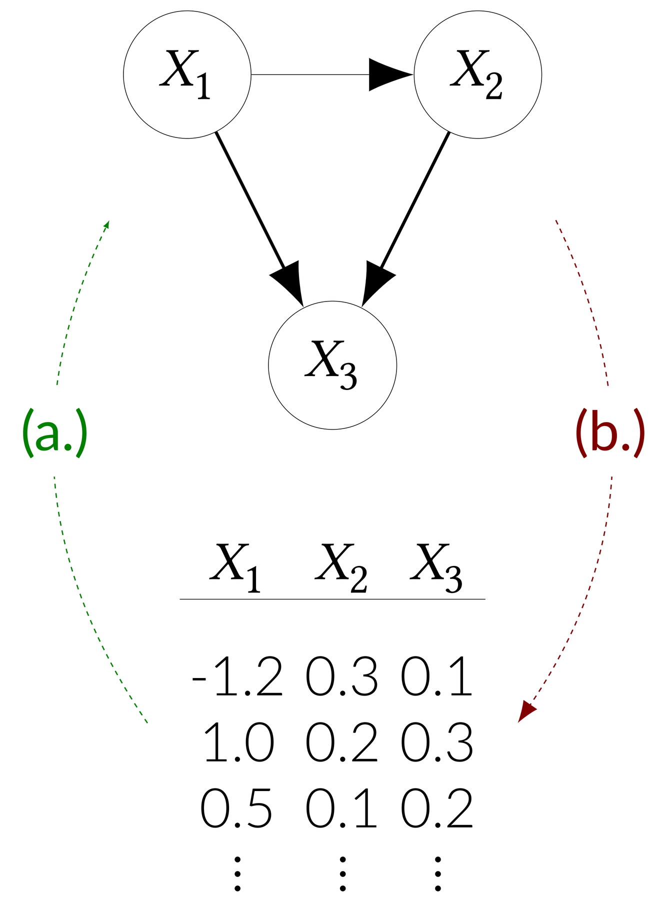


Causal Discovery

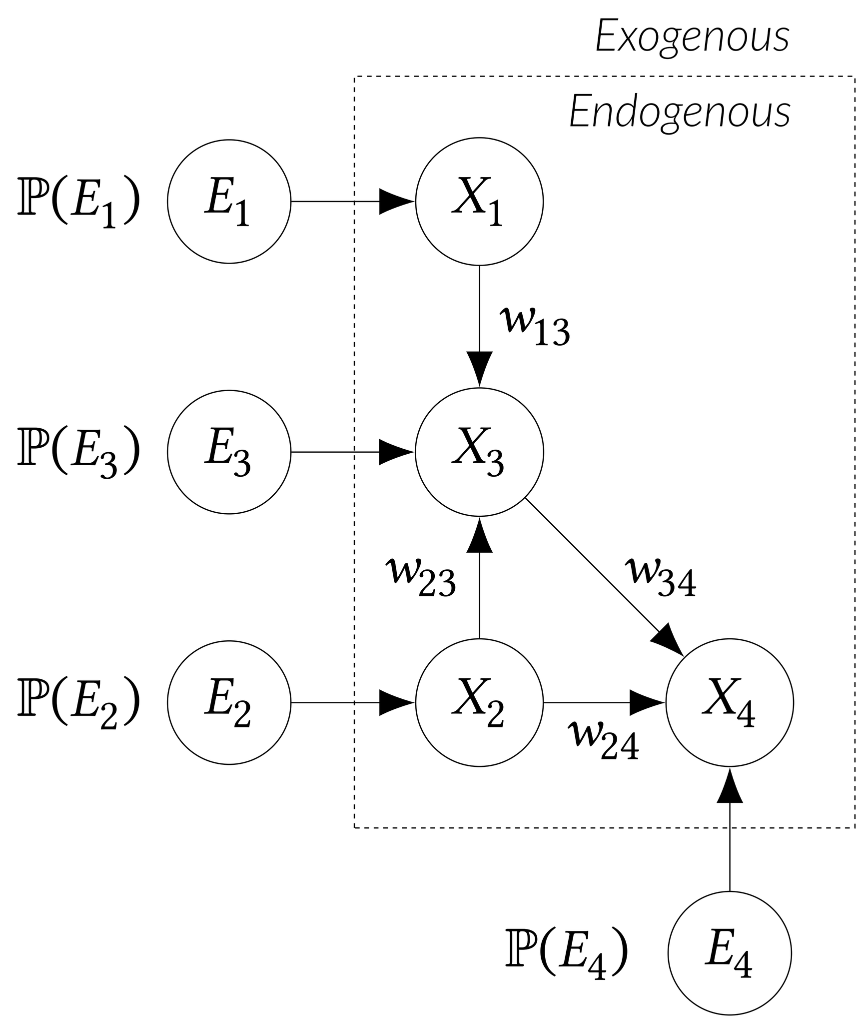
Learning causal models (a.) is challenging and generally requires non-observational data.



We can address it by restricting the **data generating process** (b.).



Structural Causal Models



Score-Based Learning

Loh and Bühlmann (2014) prove that the following score-based approach

$$\arg\min_{\mathbf{W}} \|\mathbf{W}\mathbf{X} - \mathbf{X}\|_2^2 \text{ s.t. } \mathcal{G}_{\mathcal{M}} \text{ is acyclic,}$$

has as unique minimizer the ground truth model whenever the noise distribution $\mathbb{P}_E$ is **homoscedastic**.

⚠️ ...and for **heteroscedastic** noise?

Acyclic Optimization

The space of acyclic graphs is **combinatorial**, hence expensive to search.

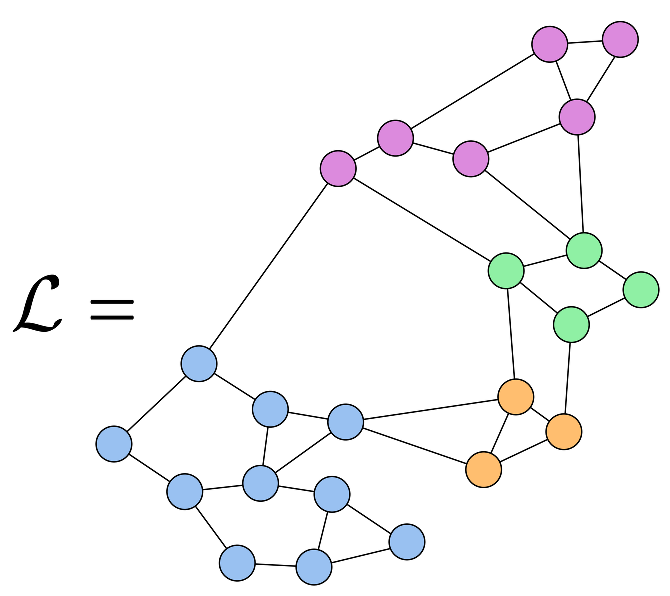
Differentiable approximations of the acyclicity constraint require $O(d^3)$ operations (Zheng et al, 2018). By losing accuracy, some methods reduce this to $O(d^2)$ (Yu et al, 2019; Massidda et al, 2024).

⚠️ Can we define **faster** methods?

⚠️ Can we handle **unobserved** data?

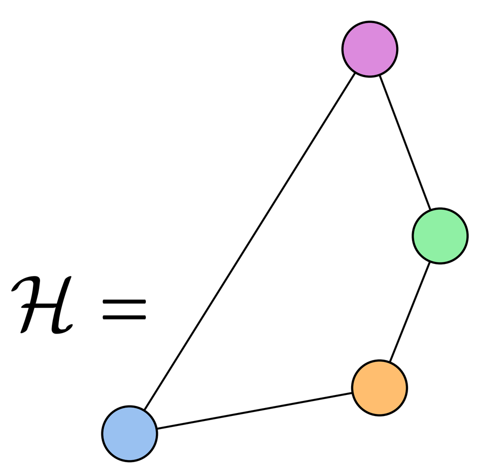
Low-Level

A **concrete** SCM represent sensor data, raw measurements, or high-dimensional data.



High-Level

An **abstract** SCM contains summary statistics, overviews, or low-dimensional representations.



Causal Abstraction

Causal Abstraction theory enables the transformation of a low-level SCM $\mathcal{L}$ into a high-level SCM $\mathcal{H}$ (Beckers et al, 2019).

Graphical and parametrical properties for **linear** SCMs are known, and they can be learned from **paired** observations (Massidda et al, 2024).

⚠️ ...and for **non-linear** SCMs?

⚠️ Can we handle **unsupervised** data?

Sounds interesting? Scan the QR!