Human-aided Discovery of Ancestral Graphs

Tiago da Silva, Eliezer da Silva, Antonio Góis,
Samuel Kaski, Dominik Heider, Diego Mesquita, Adéle Ribeiro

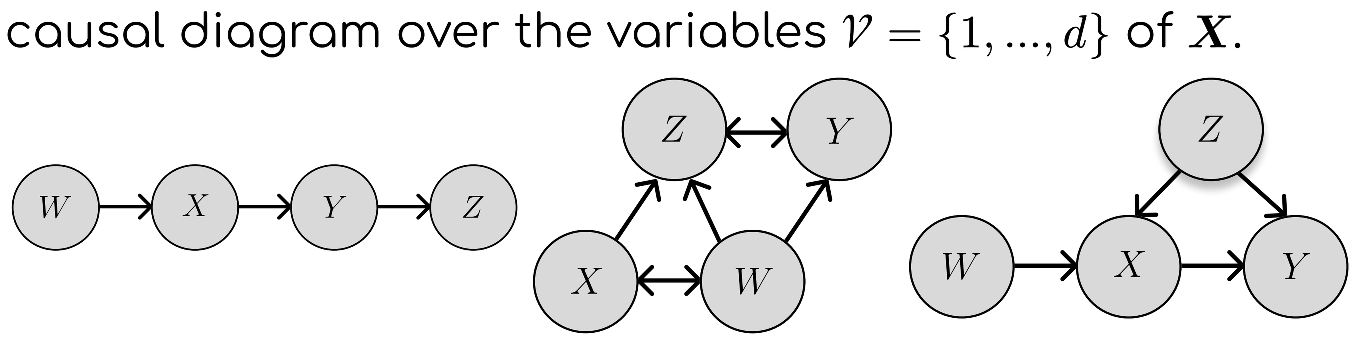
Keywords — Causal Discovery, Human in the loop, Probabilistic inference

TL;DR

- we introduce **Ancestral GFlowNets** (AGFNs) as a new amortized inference method for sampling from a belief distribution on the space of ancestral graphs,
- we develop the first human-in-the-loop framework for ancestral causal discovery (CD),
- we design an optimal strategy for elicitation of an expert's feedback regarding the nature of a specific causal relationship among the observed variables,
- we demonstrate that our human-aided CD method drastically outperforms traditional CD algorithms after just a few expert interactions.

I. Background: Causal Discovery

Let $\boldsymbol{X} \in \mathbb{R}^{n \times d}$ be a d-dimensional i.i.d. data set. A **causal discovery** (CD) algorithm takes $\boldsymbol{X}$ as input and returns a causal diagram over the variables $\mathcal{V} = \{1, \ldots, d\}$ of $\boldsymbol{X}$.



Examples of ancestral graphs.

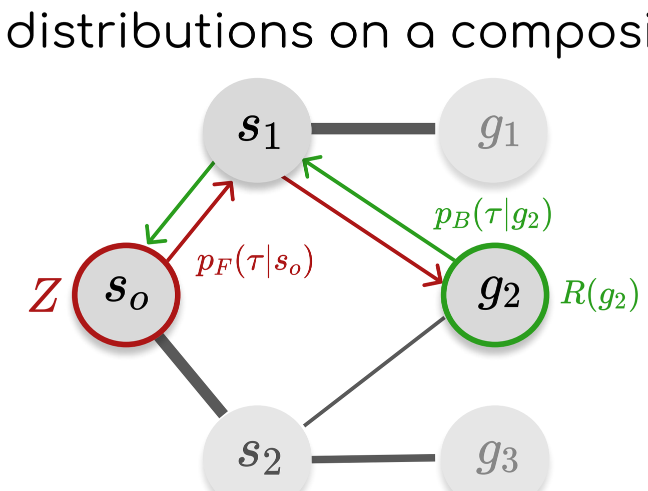
In the absence of causal sufficiency, ancestral graphs (AGs) are used to represent both ancestral causal relationships (directed edges) and associations due to latent confounding (bidirected edges) among variables.

We take a Bayesian stance and estimate a probability distribution over the space $\mathcal{G}$ of AGs on $\mathcal{V}$. For this, we introduce a score function $s : \mathbb{R}^{n \times d} \times \mathcal{G} \to \mathbb{R}$ and define the posterior distribution over the space of AGs as

$$\pi(G \mid \boldsymbol{X}) \propto \exp(s(\boldsymbol{X}, G)). \tag{1}$$

II. Background: GFlowNets

GFlowNets are amortized algorithms for sampling from unnormalized distributions on a compositional space $\mathcal{G}$.



We construct a **state graph** on the extended space $\{s_o\} \cup \mathcal{S} \cup \mathcal{G}$ endowed with an initial state s_o. Then, we learn a forward (backward) $p_F(\tau)$ ($p_B(\tau|x)$) policy s.t., for every $g \in \mathcal{G}$,

$$p_F(\tau) = \prod_{(s,s') \in \tau} p_F(s' \mid s) \text{ and } \sum_{\tau \rightsquigarrow g} p_F(\tau) \propto R(g), \tag{2}$$

in which $\tau \rightsquigarrow g$ is a trajectory starting at s_o and finishing at g. Figure 1 illustrates a state graph on $\mathcal{G} = \{g_1, g_2, g_3\}$.

To achieve this, we parameterize $p_F(\tau)$ and $p_B(\tau|x)$ as neural networks trained by stochastically minimizing

$$\mathcal{L}_{TB}(p_F, p_B) = \mathbb{E}\left[\left(\log \frac{p_F(\tau) Z}{p_B(\tau \mid x) R(x)}\right)^2\right]. \tag{3}$$

III. Ancestral Generative Flow Networks

AGFN builds upon a GFlowNet to approximate the posterior in (1); it is composed of a **state graph** and a **score function**.

1. The **state graph** (SG) is defined by an edge-addition process illustrated below. Importantly, we remove the transitions leading to non-ancestral graphs from the SG.
2. Given a model $f(\boldsymbol{X} \mid \mathcal{G}, \theta)$ indexed by parameters θ, we define the score function s as the opposite of the BIC, i.e.,

$$s(\boldsymbol{X}, G) = 2 \max_{\theta} f(\boldsymbol{X} \mid \mathcal{G}, \theta) - |E| \log n - 2\,|E| \log |V|, \tag{4}$$

in which $G = (V, E)$ and n is the size of $\boldsymbol{X}$. In this work, $f(\cdot \mid \mathcal{G}, \theta)$ is represented by a Gaussian Structural Equation Model.

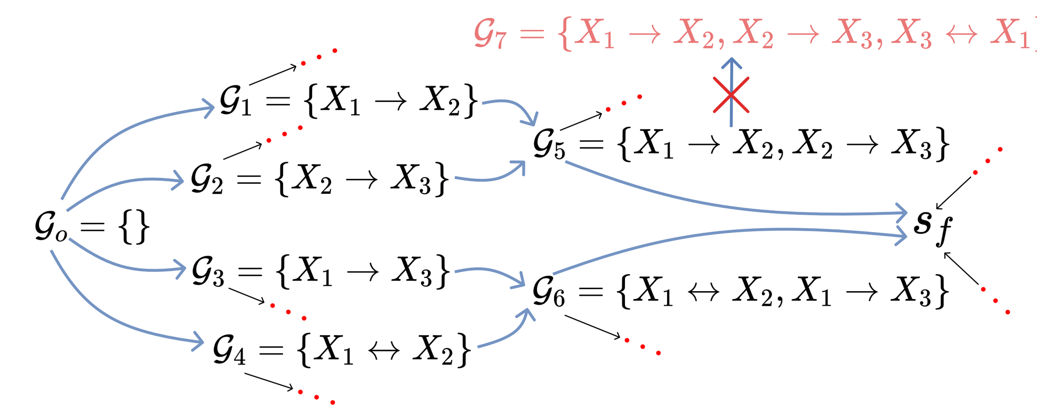


Figure 2: AGFN iteratively adds edges to an initially edgeless AG. In doing so, it ensures the sampled graphs' ancestrality.

In contrast to prior art, AGFN is strictly supported on the space of AGs. In this regard, it is the **only probabilistic method** suitable for Bayesian ancestral causal discovery.

IV. Optimal Knowledge Elicitation

Our human-in-the-loop framework has two ingredients.

1. A model of a **potentially noisy expert**: for variables V, W,

$$q\left(V \hat{\mathcal{R}} W \mid \mathcal{R}\right) = \pi \cdot 1_{V\hat{\mathcal{R}}W = V\mathcal{R}W} + \left(\frac{1-\pi}{3}\right) \cdot 1_{V\hat{\mathcal{R}}W \neq V\mathcal{R}W} \tag{5}$$

in which $\mathcal{R} \in \{\to, \leftarrow, \leftrightarrow, \emptyset\}$ ($\hat{\mathcal{R}}$) is the expert-provided (estimated) relationship between V and W; π is an hyperparameter.

2. A **scheme for integrating the expert's knowledge** into AGFN's learned model. Given feedbacks $\mathcal{F} = \{V_i \mathcal{R}_i W_i\}_{i=1}^{n}$,

$$p(G \mid \mathcal{F}) = \underbrace{p(G)}_{\text{AGFN}} \prod_{1 \le i \le n} q\left(\underbrace{V_i \mathcal{R}_i^G W_i}_{\text{Relation in G}} \mid \underbrace{V_i \mathcal{R}_i W_i}_{\text{Feedback}}\right). \tag{6}$$

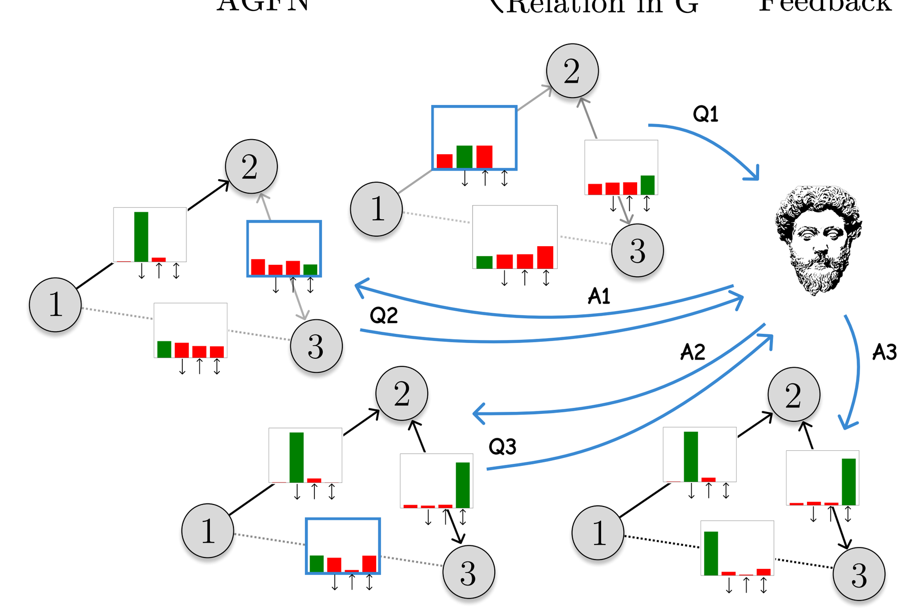
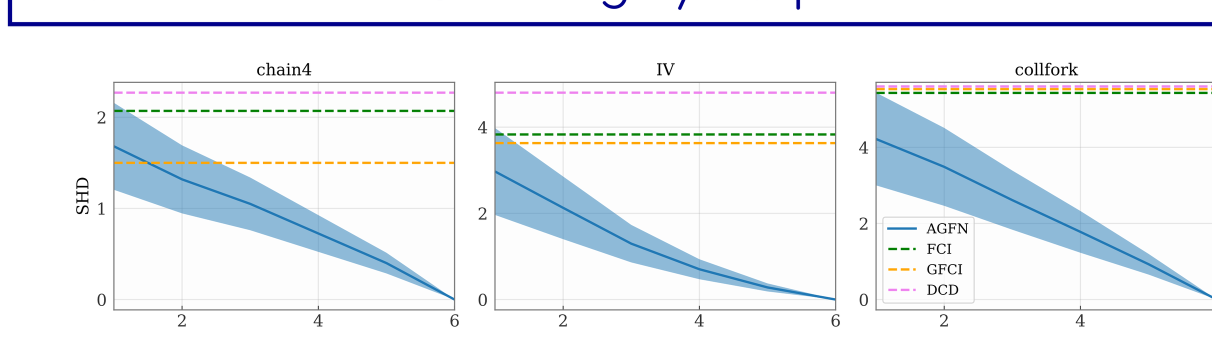


Figure 3: We progressively refine the learned AGFN through the incorporation of feedbacks from an human expert.

We probe the expert on the relation $\mathcal{R}$ minimizing the cross-entropy between distributions $p(\cdot \mid \mathcal{F} \cup \{\mathcal{R}\})$ and $p(\cdot \mid \mathcal{G})$.

V. Experimental evaluation

Human-aided AGFN largely outperforms baselines.



https://github.com/ML-FGV/agfn LatinX @ NeurIPS 2024