Amortized Causal Discovery via Prior-Fitted Networks

Complex real-world systems are often composed of simpler, sparsely interacting components. Modeling these interaction patterns enhances efficiency, interpretability, and generalization, resulting in models that are more robust to distribution shifts and local changes. Causal discovery focuses on uncovering these underlying interactions from observational or interventional data, playing a crucial role in fields like medicine, biology, economics, and climate science. By distinguishing true causal relationships from mere correlations, causal discovery provides a principled foundation for deeper scientific understanding and more effective decision-making.

Recent advances in neural network architectures have led to a new generation of causal discovery techniques, offering unprecedented scalability, flexibility, and computational efficiency. Despite these improvements, the accuracy of these methods remains constrained by statistical limitations in the likelihood estimators they employ. Additionally, margin for errors decrease rapidly with increasing graph size and density, resulting in unrealistic data volume requirements that limit practical applicability.

A promising remedy to this limitation of current causal discovery methods is the amortized approach, which aims to reuse information across multiple datasets to achieve more accurate and efficient predictions. The community has explored two main directions within this paradigm. One involves knowledge-based methods that leverage LLMs to extract domain knowledge and guide causal graph construction, while the other focuses on training predictive models to infer causal structure directly from data. However, these approaches heavily rely on either LLMs resoning skill and its understanding of the causal variables or large amounts of labeled synthetic data, which often differ significantly from real-world datasets, limiting their generalizability and practical usefulness.

In parallel, prior-fitted networks have emerged as a promising paradigm, incorporating domain-specific knowledge through specialized pretraining to enhance test-time performance. These models have demonstrated superior accuracy, especially in complex or data-scarce settings, compared to classical models trained from scratch.

Building on these insights, this project proposes a novel amortized causal discovery framework that tackles two major limitations: estimator accuracy and domain generalization. Our approach amortizes the data-dependent likelihood estimation stage by integrating domain-specific priors via pre-training on large, real-world datasets, following the principles of prior fitting. The enhanced likelihood estimators will then be employed within well-established causal structure optimization frameworks, yielding a more interpretable and statistically grounded alternative to current amortized methods.

The research will be conducted in four phases.

1. **Dataset Preparation** We will develop realistic training and evaluation datasets, with a strong emphasis on real world data.

2. **Prior-Fitting Exploration** We will evaluate various prior fitting techniques in the context of causal discovery and likelihood estimation.

3. **Model Integration and Benchmarking** We will integrate our improved estimators into a causal discovery pipeline and benchmark their performance against existing amortized and nonamortized approaches.

4. **Scalability Investigation** We will assess the scalability of the proposed methods on large datasets and explore strategies to enhance efficiency.

This work has the potential to significantly advance the field of causal discovery by delivering models that are not only fast and accurate, but also interpretable and theoretically well justified. Furthermore, our dataset preparation effort will have infrastructural value, supporting the broader research community.