Research Statement

Burak Varıcı

I am currently a Ph.D. candidate at ECSE, Rensselaer Polytechnic Institute (RPI), under the guidance of Prof. Ali Tajer, and I will obtain my degree by the Summer of 2024. I have been selected a 2020 IBM AI Horizons Scholar as part of the RPI-IBM AIRC program and worked closely with several IBM researchers, including Dr. Karthikeyan Shanmugam and Dr. Prasanna Sattigeri. **My primary research interests involve developing methodology for using interventional data to enable scalable applications of causality.** My motivation stems from a fundamental challenge in natural and applied sciences: the *identification of cause-effect relationships*. This problem entails distinguishing causation from correlation and inferring the direction of causation – often an insurmountable task with only independent and identically distributed (i.i.d.) observational data. However, interventional data hold the potential to differentiate between various causal explanations aligned with a given observation. Furthermore, recent developments in various applications have facilitated access to interventional data, offering crucial insights for answering causal questions, such as designing an optimal sequence of experiments to maximize healthcare utility or identifying unobserved causal variables.

In my doctoral research, I have leveraged the variances/invariances of causal mechanisms across different data environments. Specifically, I used interventions to uncover the underlying knowledge of causal systems by analyzing the differences caused by these interventions. I have applied my methodology to a range of causal learning problems, resulting in publications in flagship venues of AI/ML. In the following, I provide a concise overview of my contributions to the field and discuss my future research agenda.

1 Research Achievements

1.1 Intervention Target Estimation (NeurIPS'21 [1], UAI'22 [2])

Interventions in a system can manifest in two ways: (i) deliberate targeting of specific components by an experimenter, or (ii) unintended changes occurring within the system – such as the root causes of failure in cloud systems. **In my research, I tackle the challenge of estimating intervention targets in linear structural equation models.** My key idea involves hierarchically scrutinizing subsets of causal variables and investigating the invariances of their statistical features. Specifically, I devised algorithms with theoretical guarantees that not only identify intervention sites but also address two crucial gaps in the literature: (i) sample complexity analysis, particularly valuable when interventional data are limited, and (ii) scalability, achieving at least an order of magnitude improvement in computational complexity. This enhancement allows the application of interventional causal discovery algorithms beyond a trivial number of variables. Importantly, my studies cover both causally sufficient and insufficient models with latent confounders.

1.2 Causal Bandits for Intervention Design (JMLR'23 [3])

The causal bandit framework builds on ideas from the experimental design literature, integrating concepts from bandits and causality. In my research, **I aim to design algorithms that select an optimal sequence of interventions to maximize a pre-defined utility of the causal variables,** for example, identifying the best treatment option to improve a patient's health condition. To achieve this goal, I proposed algorithms that attain optimal regret scaling for causal systems with a linear structural equation model. Existing studies in the literature typically accomplish this task by using distributional information for every possible intervention, which is often a prohibitive requirement. In contrast, my algorithm dispenses with this assumption by sharing information between all interventions after each step, simultaneously improving the accuracy of their estimates. I also established previously unknown lower bounds for the problem and showed that the performance guarantees of the proposed algorithms match the scaling behavior of the lower bounds.

1.3 Causal Representation Learning ([4, 5])

Causal variables are not always observed directly but can be latent, only observed through an unknown transformation function. Causal representation learning (CRL) aims to invert the data generation process of high-dimensional observations and recover the underlying causal variables and their relationships. **In my research, I develop algorithms that are provably correct and uniquely recover latent causal variables and their causal graph.** Specifically, I proposed a novel framework (Score-based CRL) that leverages interventional data and properties of the score functions (gradient of log-likelihood) to identify latent causal variables and their relationships. My work established the first identifiability results for nonparametric causal models under a linear transform [4], and the first provably correct algorithm for the setting of general nonparametric transformations while also removing faithfulness assumptions on causal models [5].

2 Research Agenda

2.1 Causal Representation Learning: Theory meets Practice

In my Ph.D., I have contributed to the field by developing provably correct algorithms through constructive theoretical results. However, there remains a gap between existing theoretical insights and practical applications. Moving forward, I aim to bridge this gap in two intertwined directions: (i) developing theoretical results under less stringent assumptions, and (ii) applying my methodology to real-world problems.

Multi-target interventions with non-parametric settings: Existing studies focus on single-node interventions, altering only one causal mechanism in each interventional dataset. However, meeting this requirement can be challenging, e.g., interactions in robotics are likely to affect multiple causal variables at a time. To understand the applicable data settings for CRL, I plan to investigate the necessary and sufficient conditions for a set of interventional datasets to enable unique recovery of underlying causal variables.

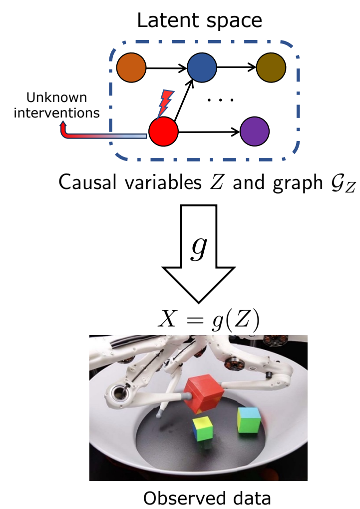

Figure 1: Causal representation learning

Universal algorithms with context-dependent theoretical guarantees: CRL algorithms are often tailored to specific assumptions on the latent model and transform. However, practical desirability calls for algorithms working in various settings. I am currently developing a score-based CRL algorithm that is agnostic to the intervention type and latent causal model, offering varying levels of theoretical guarantees, such as weaker results for soft interventions. This aligns with the main motivation of CRL, which is obtaining a useful causal representation – not necessarily a unique one. In my postdoctoral research, I plan to extend my findings and develop universal CRL algorithms for fully non-parametric systems.

Finite-sample analysis and intervention extrapolation in CRL: Another missing component of existing CRL algorithms is finite-sample analysis. Probabilistic identifiability results for a given number of interventional data samples can be useful, especially in applications where performing interventions is costly. Furthermore, certain interventions can be infeasible or unethical. Hence, it becomes crucial to infer (i) partial causal information from a small number of samples, and (ii) knowledge about unseen interventions from existing interventions. To that end, I plan to investigate the sample complexity analysis and intervention extrapolation in my score-based CRL framework.

CRL and LLMs. LLMs are trained to learn useful representations from large amounts of unlabeled data. Ideally, these high-level representations should capture the essential features of the data by focusing on the underlying process so that the models can generalize well to unseen environments. However, the current paradigm largely focuses on statistical correlations and fails to perform reliably, especially in causal reasoning tasks. The ambitious goal of CRL is learning representations that support interventions, reasoning, and planning. Motivated by the similarity of high-level objectives, I aim to leverage the ideas to solve CRL, e.g., using invariance of causal mechanisms to guide the inference process, in helping LLMs learn more robust representations that support causal reasoning. As the starting point, I am interested in investigating CRL from a mixture of interventional datasets, so that we can develop insights into the feasibility of learning causal representations from unclustered combinations of datasets, similar to the training corpus of LLMs.

2.2 Causal Discovery in Mixture of DAGs

In various applications, the observed data is heterogeneous, meaning that all causal relationships cannot be explained via a single Directed Acyclic Graph (DAG). This heterogeneity is particularly relevant to physical applications where the underlying system is constantly changing which makes it challenging to cluster the data. **In my research, I aim to develop theoretical foundations for representing such heterogeneous data using a mixture of DAGs and algorithms for learning the causal relationships within the data.**

I have started investigating this problem by formalizing the *emergent edges* – spurious connections between two causal variables that arise only due to the mixing of different DAGs. These connections cannot be differentiated from the true causal relationships without additional assumptions. Despite its practical relevance, this specific problem has not been addressed in existing literature. As a simplified version of the problem, I have established the necessary and sufficient conditions for a pair of causal variables to form an emergent edge in a mixture of polytrees [6]. I also proposed a novel graphical characterization to represent a mixture of DAGs and an algorithm for learning the discoverable causal relationships. In my future research, I plan to extend my methodology to cover a mixture of general DAGs and investigate the causal discovery of such mixtures via interventions with (partial) known information.

Related Publications

[1] B. Varici, K. Shanmugam, P. Sattigeri, and A. Tajer, "Scalable intervention target estimation in linear models," in *Proc. Advances in Neural Information Processing Systems*, pp. 1494–1505, December 2021.

[2] B. Varici, K. Shanmugam, P. Sattigeri, and A. Tajer, "Intervention target estimation in the presence of latent variables," in *Proc. Conference on Uncertainty in Artificial Intelligence*, (Eindhoven, Netherlands), pp. 2013–2023, August 2022.

[3] B. Varici, K. Shanmugam, P. Sattigeri, and A. Tajer, "Causal bandits for linear structural equation models," *Journal of Machine Learning Research*, vol. 24, no. 297, pp. 1–59, 2023.

[4] B. Varıcı, E. Acartürk, K. Shanmugam, A. Kumar, and A. Tajer, "Score-based causal representation learning with interventions," *arXiv:2301.08230*, 2023.

[5] B. Varıcı, E. Acartürk, K. Shanmugam, and A. Tajer, "General identifiability and achievability for causal representation learning," *arXiv:2310.15450 (under review for AISTATS)*, 2023.

[6] B. Varıcı, D. Katz, A. Tajer, D. Wei, and P. Sattigeri, "Separability analysis for causal discovery in mixture of dags," *accepted for publication in Transactions on Machine Learning Research*, 2024.