CORE: Towards Scalable and Efficient Causal Discovery with Reinforcement Learning

Anonymous CVPR submission

Paper ID *****

Abstract

Causal discovery (CD) is the challenging task of inferring causal connections between a set of variables from data [4, 14]. Most traditional approaches to CD consider data from pure observations of the investigated scenario. These are approaches such as constraint-based ones [6, 12], score-based ones [3], and more recently continuous optimization-based ones [15, 16].

In the context of CD, Pearl's Causal Hierarchy (PCH) asserts that distinguishing between mere correlations and genuine causal relationships requires considering interventions in general [1]. As a response to this requirement, there has been a recent push to incorporate interventions into causal discovery research [5, 9] including machine learning [2, 7, 11], among others.

Reinforcement learning (RL) learns an optimal policy for sequential decision problems through interactions [13]. Therefore, RL is a promising framework for using interventions to investigate causal relationships by framing CD as a sequential decision problem. In particular, RL plays a dual role in the realm of causal discovery - it can be used not only to recover the causal structure of an environment [17], but also to learn causal discovery algorithms [11], thus representing a versatile tool for CD.

Although causal discovery has seen substantial progress, challenges persist in areas such as scalability, generalization, and planning of interventions. In this context, this paper introduces CORE (**C**ausal Disc**O**very with **RE**inforcement Learning), a deep-RL-based [10] approach designed for the task of learning a CD algorithm. CORE learns a policy that sequentially reconstructs causal graphs from both observational and interventional data, while simultaneously performing informative interventions. By providing a limited horizon, these interventions and the number of samples are highly budgeted, which is desirable in CD in general. The dual learning paradigm allows CORE not only to uncover causal structures efficiently, but also to identify interventions that enhance its causal models. The following lists our main contributions:

- We formalize the task of learning a CD algorithm as a partially observable Markov decision process (POMDP).
- We propose a dual Q-learning setup to simultaneously learn intervention design and structure estimation more efficiently.
- We demonstrate that CORE can be successfully applied for causal discovery to previously unseen graphs of sizes of up to 10 variables.

In addition, we show the importance of jointly learning which interventions to perform and graph generation and investigate the limitations of our approach regarding its applicability to the real world.

The most distinctive feature of CORE is that it does not impose a specific algorithm for identifying causal models, but rather attempts to learn it. This can have positive effects on efficiency and transferability to new problem instances. While MCD [11] and AVICI [8] solve the same task, they run into pitfalls that hinder their application to realistic graph sizes or rely on offline data, respectively. We set steps to overcome these pitfalls by imposing additional structure on our policy, more efficient rewards, and learning to actively perform relevant interventions.

Our results show robust generalization to unseen graphs and the capability to scale to scenarios with up to ten variables, a step forward over the state of the art of learning CD algorithms, and a crucial advancement towards addressing real-world complexities. The joint learning of intervention selection and graph generation is shown to be crucial.

Overall, CORE provides an automated and adaptable approach for uncovering causal relationships in complex systems. This aligns well with the goals of object-centric representations and causal reasoning for robotics, where robust and efficient structure learning from interaction data is key. Although limited, the scalability demonstrated is relevant for some real-world robotic domains. This work was recently accepted at AAMAS 24 and code is available at https://anonymous.4open.science/r/CORE-C0BF/. We believe it makes a strong contribution towards the workshop's aim of advancing causal learning for embodied AI.

CVPR
#*****

References

[1] Elias Bareinboim, Juan D. Correa, Duligur Ibeling, and Thomas Icard. On Pearl's Hierarchy and the Foundations of Causal Inference. In *Probabilistic and Causal Inference*, pages 507–556. ACM, New York, NY, USA, 2022. 1

[2] Philippe Brouillard, Sébastien Lachapelle, Alexandre Lacoste, Simon Lacoste-Julien, and Alexandre Drouin. Differentiable Causal Discovery from Interventional Data. In *Advances in Neural Information Processing Systems*, pages 21865–21877. Curran Associates, Inc., 2020. 1

[3] David Maxwell Chickering. Optimal Structure Identification with Greedy Search. *J. Mach. Learn. Res.*, 3(null):507–554, 2003. 1

[4] Clark Glymour, Kun Zhang, and Peter Spirtes. Review of Causal Discovery Methods Based on Graphical Models. *Frontiers in Genetics*, 10, 2019. 1

[5] Alain Hauser and Buhlmann@stat Math Ethz Ch. Characterization and Greedy Learning of Interventional Markov Equivalence Classes of Directed Acyclic Graphs Peter B ¨ uhlmann. *Journal of Machine Learning Research*, 13:2409–2464, 2012. 1

[6] Antti Hyttinen, Frederick Eberhardt, and Matti Järvisalo. Constraint-based Causal Discovery: Conflict Resolution with Answer Set Programming. In *Conference on Uncertainty in Artificial Intelligence*, pages 340–349, 2014. 1

[7] Phillip Lippe, Taco Cohen, and Efstratios Gavves. Efficient Neural Causal Discovery without Acyclicity Constraints. In *International Conference on Learning Representations*, 2021. 1

[8] Lars Lorch, Scott Sussex, Jonas Rothfuss, Andreas Krause, and Bernhard Schölkopf. Amortized Inference for Causal Structure Learning. In *Advances in Neural Information Processing Systems*, pages 13104–13118. Curran Associates, Inc., 2022. 1

[9] Joris M Mooij, Sara Magliacane, and Tom Claassen. Joint Causal Inference from Multiple Contexts. *Journal of Machine Learning Research*, 21:1–108, 2020. 1

[10] Aske Plaat. *Deep reinforcement learning*. Springer, 2022. 1

[11] Andreas W M Sauter, Erman Acar, and Vincent Francois-Lavet. A Meta-Reinforcement Learning Algorithm for Causal Discovery. In *Proceedings of the Second Conference on Causal Learning and Reasoning*, pages 602–619. PMLR, 2023. 1

[12] Peter Spirtes, Clark N Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2000. 1

[13] R.S. Sutton and A.G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018. 1

[14] Matthew J. Vowels, Necati Cihan Camgoz, and Richard Bowden. D'ya Like DAGs? A Survey on Structure Learning and Causal Discovery. *ACM Computing Surveys*, 55(4): 1–36, 2023. 1

[15] Yue Yu, Jie Chen, Tian Gao, and Mo Yu. DAG-GNN: DAG Structure Learning with Graph Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning*, pages 7154–7163. PMLR, 2019. 1

[16] Xun Zheng, Bryon Aragam, Pradeep Ravikumar, and Eric P Xing. DAGs with NO TEARS: Continuous Optimization for Structure Learning. 2018. 1

[17] Shengyu Zhu, Ignavier Ng, and Zhitang Chen. Causal Discovery with Reinforcement Learning. 2019. 1