Project Proposal: Discover Causality by Exploiting the Structural Causal Model (SCM) by Reinforcement Learning Agent

Jincheng Zhou **Yang Li** **Jiaxin Cheng** **Tianmu Lei**
University of Southern California
{jinchenz, yli546, jiaxinch, tianmule}@usc.edu

1 Motivation

[1] clarifies the learning process of a causal DAG given data points during learning from a Probabilistic Graphical Model point of view (Bayesian Network learning with Dirichlet distribution).

[2] on the other hand, seems like some work that's "downstream" to our work in that our result (learning a causal model) can be combined with that one's (exploiting a causal model for guided policy search) and consequently be generalized to a new model-based reinforcement learning framework (one that both learns and exploits causal models).

Therefore, it leads us to thinking how we can approach the problem of learning discrete and concrete state embeddings. i.e., we need the agent to have a clear mind of knowing that certain states or events correspond to certain nodes in the causal model, so that it can gradually construct the causal model in the first place before even learning it.

Basically we are looking for some mechanism that endows the agent with the capability of learning "state encodings", one that probably comes with certain semantic meanings. Theoretically we can then use this finite set of state encodings of the environment as the finite set of nodes in the causal graph.

2 Method

- GCN and RNN
- (meta) RL: DQN, Policy Gradient, PPO, Actor-Critic, A2C, A3C, etc.

3 Problem Formulation

- Input: Causal Bayesian Network (CBN) as the innate physical mechanism of the given environment
- Output: The graph learned and encoded by the GCN agent.

4 Milestones

- Building the infrastructure for iterating models and conducting experiments for the project
- Implement and realize the results of the paper "Causal Reasoning from Meta-reinforcement Learning"
- Plug in GCN and show that GCN can learn, encode, and express toy CBNs.

32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

- Conduct thorough and rigorous experiments to show that our models and results are actually valid
- Enlarge the scale of toy CBNs and show that our models can potentially learn meaningful results in a real-world seething

5 Expected approach and results

References

[1] M. Gonzalez-Soto, L. E. Sucar, and H. J. Escalante. Playing against Nature: causal discovery for decision making under uncertainty. *arXiv e-prints*, page arXiv:1807.01268, Jul 2018.

[2] Lars Buesing, Theophane Weber, Yori Zwols, Nicolas Heess, Sebastien Racaniere, Arthur Guez, and Jean-Baptiste Lespiau. Woulda, coulda, shoulda: Counterfactually-guided policy search. In *International Conference on Learning Representations*, 2019.