

## Counterfactual Data Augmentation using Locally Factored Dynamics

Various dynamic processes in robot control consist of subprocesses which interact with each other during the execution of the global process. Sparse interactions between subprocesses allows the formulation of locally independent mechanisms for globally optimal behavior. The work proposes a novel model-free data augmentation framework by conditioning on locally factored dynamics which aid in the generation of causally valid counterfactual experiences. Counterfactual Data Augmentation (CoDA) leverages causal independence in interactions to factor local dynamics in an object-oriented manner. Improved performance of model-free agents trained on counterfactual experiences validates the suitability of CoDA in batch-constrained and goal-conditioned settings.

CoDA makes use of sparse interactions between objects in the environment to locally factor the dynamics. Given a graphical structure of environment interactions, object entities may be considered locally independent if the interactions between them are limited over successive timesteps. This leads to a decomposition of state and action spaces which is based on the minimality assumption of the environment as a factored MDP. Connected edges in the factored MDP depict local interactions which can causally model new transitions in independent subspaces. CoDA is implemented as a function of two factual transitions and mask function which yield a causal transition in the same local subspace as the original transitions. This bound is validated by making use of structured equations in local neighborhood. Practical implementation of CoDA makes use of attention masks in a transformer setting. The framework is trained to infer local factorization instead of sampling future states as in model-based RL.

Contrary to typical model-free frameworks, CoDA aims to improve data-efficiency in RL by leveraging aspects of the environment dynamics. The problem of modelling local independence in causal structures is an interesting and novel direction which depicts improvement in comparison to learned models of the environment. However, the method presents two aspects with room for improvement. Firstly, the learned variant of CoDA lags behind the variant with access to ground truth. This indicates that the model formed by counterfactual experiences may not be a complete and accurate representation of practical scenarios. A more rigorous validation on real-world robot tasks would aid in better evaluation of the proposed scheme. Secondly, increasing the amount of state factors in batch RL only leads to little increment in success rates. This depicts a lack of local dynamic information contained in the factorization scheme. A potential improvement could consist of learning factorizations based on information theoretic measures such as maximization of mutual information between counterfactual and local dynamics.

Local factorization of dynamics paves new directions for future work in regard to learning and better reasoning. As mentioned in the work, learning of the masking function and prioritization of CoDA scheme to relevant counterfactual samples is an interesting direction for future work. Additionally, the framework provides insights towards the unification of model-based and model-free methods for robust control. This would aid in better generalization to tasks where the dynamics are disentangled or sparsely accessible.

Sample-efficiency in model-free RL is an open problem for learning informative dynamics of the environment. The work aims to bridge this gap by locally factoring dynamics based on sparse interactions between objects in the environment. Local factorization assumes conditional independence over a fixed time-scale in the global process. This allows for a causal combination of counterfactual experiences which lie in the same local sub-space. CoDA agents benefit from addi-

## Review-10

---

tional augmented data and performance gains. The framework lays out directions for learning local factorizations of dynamics and their extensions to disentangled real-world scenarios.