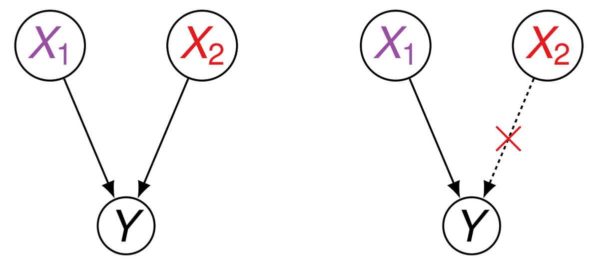
Fine-Grained Causal Dynamics Learning with Quantization for Improving Robustness in Reinforcement Learning

Inwoo Hwang Yunhyeok Kwak Suhyung Choi Byoung-Tak Zhang† Sanghack Lee†

SEOUL NATIONAL UNIVERSITY · AIIS Artificial Intelligence Institute, Seoul National University · ICML International Conference On Machine Learning

Motivation & Background

- **Causal dynamics learning** aims to build a dynamics model that makes predictions based on the causal relationships among the environmental entities.
- However, causal connections often manifest only under certain **contexts** and existing approaches overlook such **fine-grained relationships**.
- **Local independence** (e.g., context-specific independence):

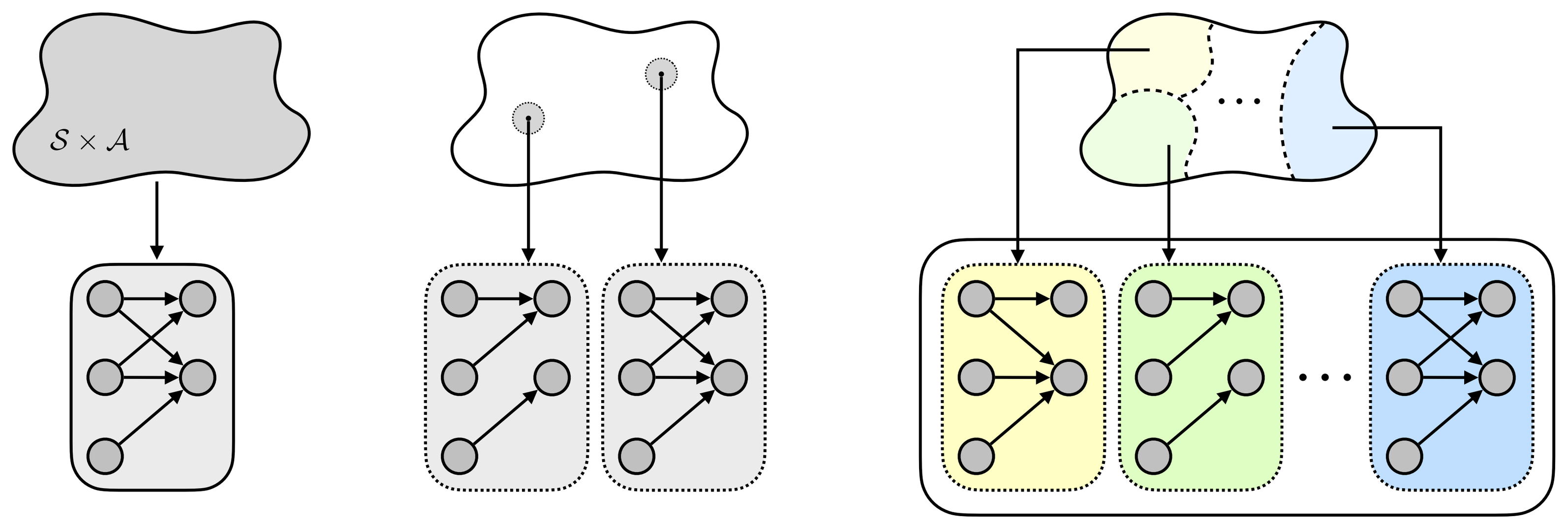


$$Y = \begin{cases} f_1(X_1, X_2, U) & \text{if} \quad (X_1, X_2) \in \mathcal{D}^c \\ f_2(X_1, U) & \text{if} \quad (X_1, X_2) \in \mathcal{D} \end{cases}$$

$\mathcal{G}$: causal graph (CG), $\mathcal{G}_\mathcal{D}$: local causal graph (LCG) on $\mathcal{D}$

  - **Physical laws**: To move a static object (Y), a force (X_2) exceeding frictional resistance (X_1) must be exerted. Otherwise, it would not move.
  - **Autonomous driving** ($Y = X_1 \vee X_2$): A car must stop ($Y = 1$) in the presence of a pedestrian on the road ($X_1 = 1$), ignoring the traffic signal (X_2).
- **Fine-grained causal reasoning** $\Rightarrow$ robustness to locally spurious correlations!

Related Works & Problem Formulation



(a) (b) (c)

- **(a)** Prior causal dynamics models: $p_\phi(s' \mid s, a; \mathcal{G})$
- **(b)** Sample-specific approaches: $p_\phi(s' \mid s, a; \mathcal{G}_{(s,a)})$
- **(c)** Our approach: $p_\phi(s' \mid s, a; \mathcal{G}_\mathcal{E})$

  ✓ fine-grained causal relationships ✓ interpretable ✓ theoretically-grounded

◎ **Goal**: Discover meaningful decomposition $\{\mathcal{E}_1, \cdots, \mathcal{E}_K\}$ of state-action space $\mathcal{S} \times \mathcal{A}$ that entails sparse LCGs $\{\mathcal{G}_1, \cdots, \mathcal{G}_K\}$ and incorporate them into dynamics modeling.

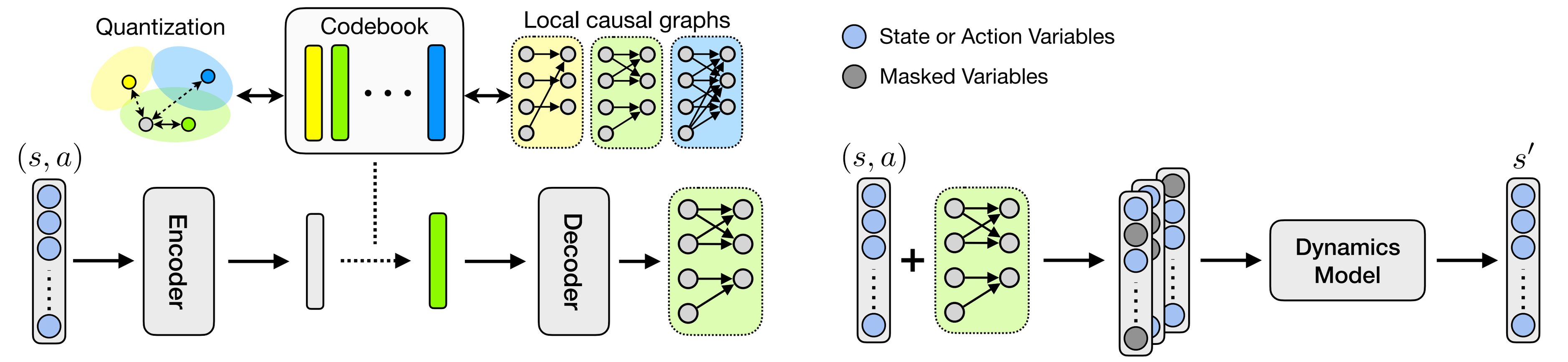
Fine-Grained Causal Dynamics Learning

Score for Decomposition and Graphs

- $\mathcal{S}(\{\mathcal{G}_z, \mathcal{E}_z\}_{z=1}^K) := \sup_\phi \mathbb{E}_{p(s,a,s')} \left[\log \hat{p}(s' \mid s, a; \mathcal{G}_z, \phi) - \lambda |\mathcal{G}_z|\right]$
- We aim to find $\{\mathcal{G}_z^*, \mathcal{E}_z^*\} \in \arg\max \mathcal{S}(\{\mathcal{G}_z, \mathcal{E}_z\}_{z=1}^K)$ with dynamics model $\hat{p}$.

Idea: Learn a discrete latent variable with **vector quantization** that represents the pairs of a subgroup $\mathcal{E}_z$ and LCG $\mathcal{G}_z$.

Overall Framework



(a) Local Causal Graph Inference (b) Masked Prediction with LCG

- **End-to-end joint training** of the dynamics model and the codebook.
- Each sample (s, a) is quantized to nearest code e_z, which is then decoded to the adjacency matrix A_z representing the graph $\mathcal{G}_z$.
- The dynamics model employs the inferred LCGs for prediction.

Training Objective

$$\mathcal{L}_{\text{total}} = \underbrace{-\log \hat{p}(s' \mid s, a; A) + \lambda \cdot \|A\|_1}_{\text{prediction loss + regularization}} + \underbrace{\|\text{sg}[h] - e\|_2^2 + \beta \cdot \|h - \text{sg}[e]\|_2^2}_{\text{quantization loss}}.$$

Theoretical Analysis & Interpretation

Theoretical Analysis

✓ Identifiability of LCGs (*Thm. 1*) ✓ Identifiability of contexts (*Thm. 2*)

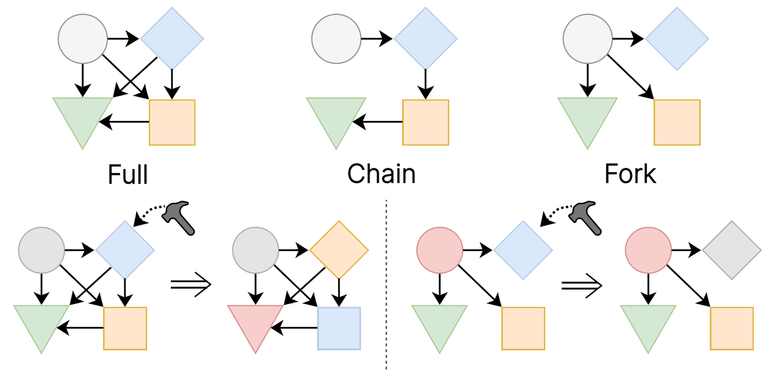
- Requirement: sufficient quantization degree K (= codebook size).

Connections to Prior Approaches

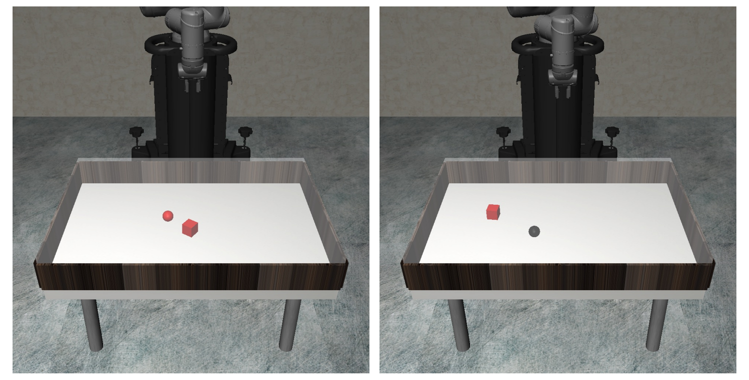
- $K = 1$: corresponds to prior (global) causal models.
- $K \to \infty$: reverts to sample-specific approaches.

Experiments

Environments



(a) Chemical



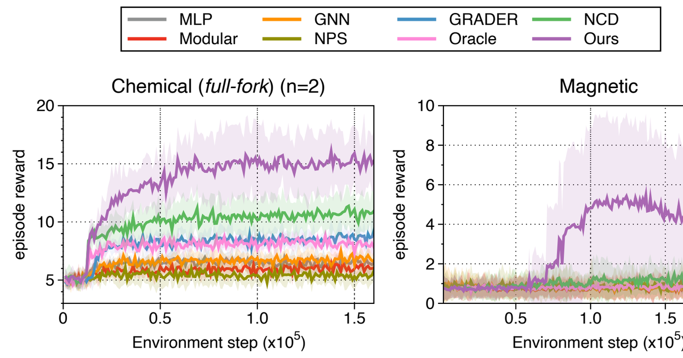
(b) Magnetic

- In Chemical, the root node determines fine-grained relationships.
- In Magnetic, an object exhibits magnetism if it is colored red.

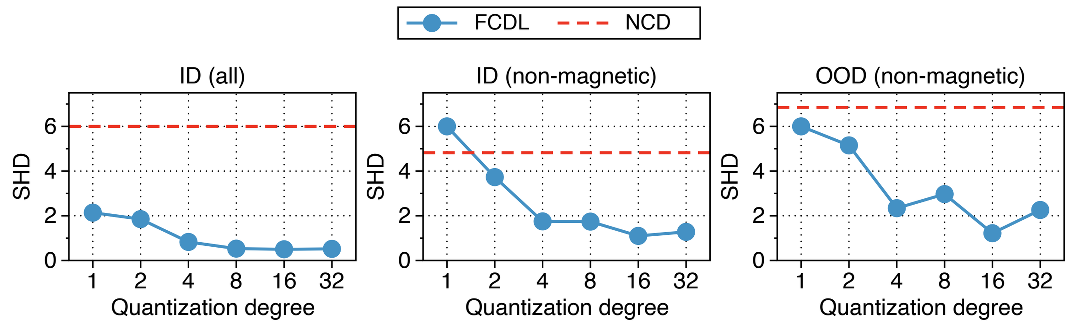
Prediction Accuracy (ID, OOD)

Setting / n		MLP	Modular	GNN	NPS	CDL	GRADER	Oracle	NCD	FCDL (Ours)
full-fork	($n = 0$)	88.31±1.58	89.24±1.52	88.81±1.44	58.34±2.08	89.22±1.67	87.75±1.64	89.63±1.62	**90.07**±1.22	89.46±1.40
	($n = 2$)	31.11±1.69	26.53±3.45	36.29±3.45	40.56±4.61	35.59±1.85	37.93±1.06	33.87±1.34	41.60±5.08	**66.44**±12.22
	($n = 4$)	30.44±2.28	24.73±5.61	25.80±3.48	26.81±4.37	35.82±1.40	38.94±1.63	36.48±1.80	37.47±2.13	**58.49**±10.20
	($n = 6$)	32.39±1.76	26.73±8.31	21.58±3.44	23.02±4.27	42.22±1.39	45.74±2.25	42.47±0.75	42.27±1.82	**49.09**±4.77
full-chain	($n = 0$)	84.38±1.31	85.92±1.15	85.41±1.84	58.48±2.81	**86.85**±1.47	84.24±1.22	85.76±1.56	85.63±1.01	86.07±1.62
	($n = 2$)	28.66±3.65	25.24±4.68	29.22±3.39	38.73±2.63	34.90±1.59	36.82±3.12	34.63±1.78	40.04±6.21	**60.34**±12.10
	($n = 4$)	26.52±4.26	24.94±4.81	23.28±4.98	27.69±4.28	36.52±1.72	37.41±2.84	38.31±2.48	37.47±2.98	**56.64**±9.40
	($n = 6$)	24.15±4.17	25.09±5.91	20.53±6.96	24.45±3.84	42.06±1.29	43.48±4.14	42.87±2.08	41.19±1.66	**53.29**±6.63

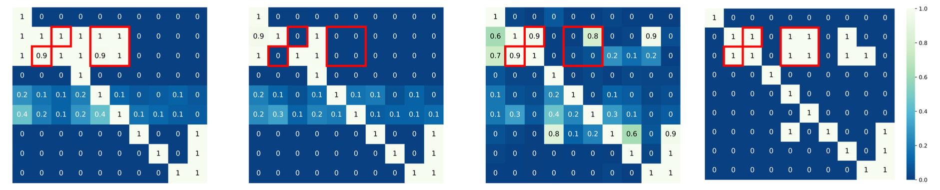
Learning Curve



Local Causal Discovery

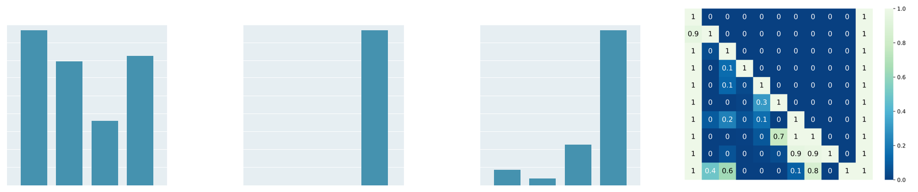


Learned LCGs



(a) FCDL (CG) (b) FCDL (LCG) (c) NCD (ID) (d) NCD (OOD)

Learned Contexts



(a) ID (all) (b) ID (*fork*) (c) OOD (*fork*) (d) Learned LCG

Conclusion

- We present a novel approach to dynamics learning that infers **fine-grained causal relationships**, leading to improved robustness of MBRL.
- Our method learns a discrete latent variable that represents the pairs of a subgroup and a local causal graph (LCG), allowing joint optimization with the dynamics model.

The 41st International Conference on Machine Learning (ICML 2024), Vienna, Austria