Causal Learning and Reasoning in Multi-Agent RL

Giovanni Briglia*

Supervisors: **Franco Zambonelli**, **Stefano Mariani**

*National PhD in AI: University of Pisa & University of Modena and Reggio Emilia, Italy

giovanni.briglia@phd.unipi.it | giovannibriglia.github.io | Giovannibriglia

DIPI Lab — DIstributed & Pervasive Intelligence Laboratory

Why causal reasoning for (MA)RL?

Single-agent RL: causal reasoning helps in special regimes — partial observability, offline data, off-policy logs.

Multi-agent RL: other agents are not part of the environment. They are **adaptive, endogenous causal mechanisms**. Non-stationarity, partial observability, and decentralisation are *causal* problems by default.

Two properties hold silently in single-agent RL, and *break* in MARL.

Observation-equivalence

$P(\cdot \mid A^i) \neq P(\cdot \mid do(A^i))$

back-door paths run through A^{-i}

Executability

$do(A^i)$ fixes *one* component of the joint action

effects depend on π^{-i} → partial, agent-relative

Research question

How to **represent**, **integrate**, and **understand** causal reasoning in MARL?

Two orthogonal dimensions for causal MARL

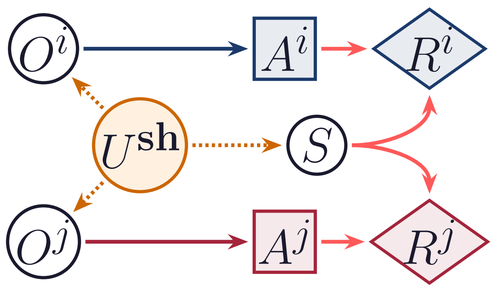
Dimension 1 — Interaction Entanglement: *why* bother with causality.

Your Q-values estimate $\mathbb{E}[R \mid S, A]$. You want $\mathbb{E}[R \mid S, do(A)]$. In RL, and mostly in MARL, these diverge — and the gap is hard to detect during training. Causal reasoning is the tool that detects, corrects, or bounds this gap.

Informational Entanglement

A latent U^{sh} drives the shared state; each O^i captures only a fragment. Conditioning ≠ intervention.

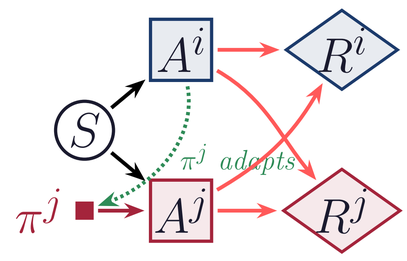
⇒ ***Identifiability fails.***



Strategic Entanglement

Full observability, but π^j adapts to π^i over time. Effects of $do(A^i)$ are response-conditioned through π^j.

⇒ ***Effects depend on π^{-i}.***



neither ⇒ classical RL suffices | **one** ⇒ causal reasoning helps | **both** ⇒ causal modeling is unavoidable

Dimension 2 — Objective Alignment: *what* causality gives you.

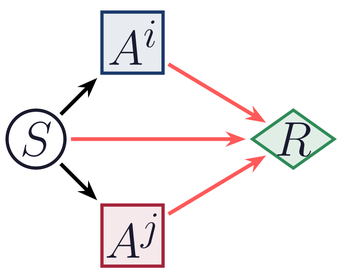
Cooperative — counterfactual credit assignment, cleaner than TD when teammates confound your actions

Competitive — strategic counterfactuals, anticipating how opponents respond to a committed policy

Aligned Objectives

Shared reward R couples both actions. The team learning signal must attribute R between A^i and A^j.

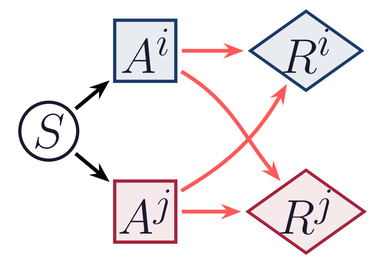
⇒ ***Coordination + counterfactual credit assignment.***



Mixed / Adversarial

Diverging rewards R^i, R^j. Each action affects both — but with opposing or partial incentives.

⇒ ***Strategic anticipation + opponent counterfactuals.***

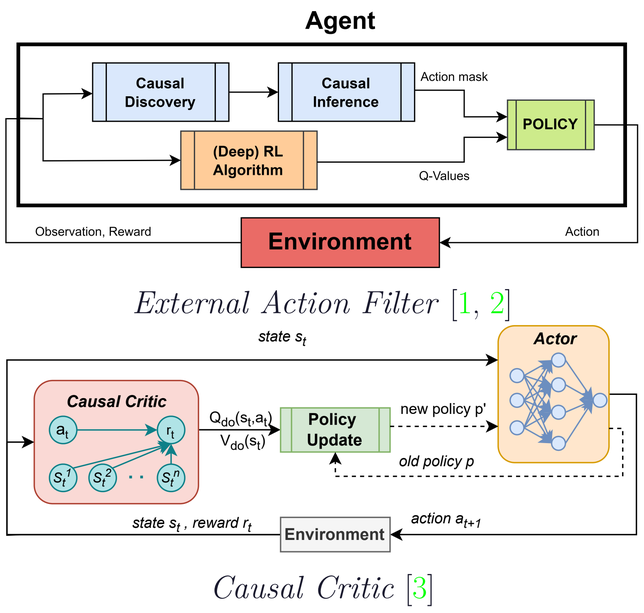


Legend: transition | reward | policy i | policy j | policy adaptation | latent | confounder

Preliminary and On-Going Results

Integration

Causal modules in the RL loop

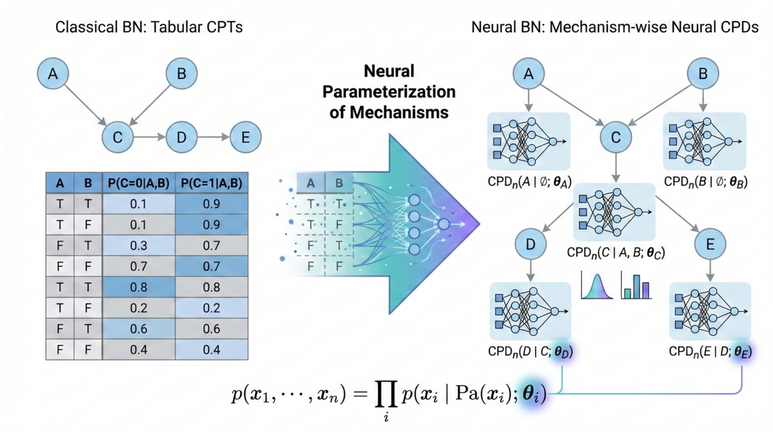


External Action Filter [1, 2]

Causal Critic [3]

Representation

Per-agent reasoning under uncertainty



Neural Bayesian Networks

- differentiable, local causal models
- inference engine behind the action filter and the critic
- tractable interventional and counterfactual queries

Foundations

- **Taxonomy** causal (MA)RL [4]
- **Roadmap** [5]
- **Formaliszation** of entanglement & alignment

Feedback wanted

- Algorithms, benchmarks and evaluation metrics for causal MARL
- Distributed causal discovery in MARL
- Understanding (MA)RL generalization through causal reasoning
- Novel trends

Selected Publications

[1] Giovanni Briglia et al. "Improving Reinforcement Learning-Based Autonomous Agents with Causal Models". In: *25th International Conference on Principles and Practice of Multi-Agent Systems PRIMA*. Vol. 15395. Lecture Notes in Computer Science. Springer, 2024, pp. 267–283.

[2] Giovanni Briglia, Stefano Mariani, and Franco Zambonelli. "Towards Safe Action Policies in Multi-robot Systems with Causal Reinforcement Learning". In: *Agents and Robots for reliable Engineered Autonomy*. Springer Nature Switzerland, 2025, pp. 51–71.

[3] Giovanni Briglia, Stefano Mariani, and Franco Zambonelli. "Causal Models Improve Reinforcement Learning for Pervasive and Robotic Tasks". In: *2026 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. In press. IEEE Computer Society. 2026.

[4] Stefano Mariani et al. "Causal Learning and Reasoning in Multi-Agent Systems: Goals, Issues, and Taxonomy". In: *Proceedings of the 1st International Workshop on Causal Learning and Reasoning in Agents and Multiagent Systems, CLaRAMAS - Paphos, Cyprus, May 26, 2026*. In press. Springer, 2026. URL: https://openreview.net/pdf?id=FCr9YFHvwO.

[5] Giovanni Briglia, Stefano Mariani, and Franco Zambonelli. "A Roadmap Towards Improving Multi-Agent Reinforcement Learning With Causal Discovery And Inference". In: *arXiv preprint arXiv:2503.17803* (2025).