

Centralized Training with Hybrid Execution in Multi-Agent Reinforcement Learning via Predictive Observation Imputation (Abstract Reprint)

Pedro P. Santos^{1,2}, Diogo S. Carvalho^{1,2}, Miguel Vasco³, Alberto Sardinha⁴, Pedro A. Santos^{1,2}, Ana Paiva^{1,2}, Francisco S. Melo^{1,2}

¹Artificial Intelligence for People and Society (GAIPS), INESC-ID

²Instituto Superior Técnico, University of Lisbon

³KTH Royal Institute of Technology

⁴Pontifical Catholic University of Rio de Janeiro

Abstract Reprint. This is an abstract reprint of the journal article by Santos, Carvalho, Vasco, Sardinha, Santos, Paiva, and Melo (2025).

Abstract

We study hybrid execution in multi-agent reinforcement learning (MARL), a paradigm where agents aim to complete cooperative tasks with arbitrary communication levels at execution time by taking advantage of information-sharing among the agents. Under hybrid execution, the communication level can range from a setting in which no communication is allowed between agents (fully decentralized), to a setting featuring full communication (fully centralized), but the agents do not know beforehand which communication level they will encounter at execution time. We contribute MARO, an approach that makes use of an auto-regressive predictive model, trained in a centralized manner, to estimate missing agents' observations at execution time. We evaluate MARO on standard scenarios and extensions of previous benchmarks tailored to emphasize the impact of partial observability in MARL. Experimental results show that our method consistently outperforms relevant baselines, allowing agents to act with faulty communication while successfully exploiting shared information.

References

Santos, P. P.; Carvalho, D. S.; Vasco, M.; Sardinha, A.; Santos, P. A.; Paiva, A.; and Melo, F. S. 2025. Centralized training with hybrid execution in multi-agent reinforcement learning via predictive observation imputation. *Artificial Intelligence*, 348: 104404.