The Fortieth AAAI Conference on Artificial Intelligence (AAAI-26)

Scaling Up Cooperative Multi-Agent Reinforcement Learning Through Hierarchical Heterogeneous Modular Architectures

Minghong Geng
School of Computing and Information Systems
Singapore Management University, Singapore
mhgeng.2021@phdcs.smu.edu.sg

Abstract

Multi-agent reinforcement learning enables sophisticated collaborative behaviors in autonomous systems, yet fundamental scalability barriers persist: existing methods struggle to coordinate large agent populations and face challenges with extended decision-making horizons. This research develops hierarchical approaches to scale up multi-agent learning systems through two complementary directions: structural scaling for coordinating increasing numbers of agents and temporal scaling for extending decision-making horizons. This paper presents four integrated contributions: a taxonomic survey establishing hierarchical architectures as the theoretical foundation for scalable multi-agent learning systems, a benchmark for long-horizon multi-objective multi-agent reinforcement learning, a framework integrating self-organizing neural networks with multiple reinforcement learning agents for hierarchical tri-level control, and a framework leveraging large language models for zero-shot multi-agent planning. Through comprehensive validation, this work demonstrates that hierarchical, heterogeneous, modular architectures provide unified, interpretable solutions to multi-agent scalability, bridging theoretical multi-agent reinforcement learning research with real-world deployment requirements.

Introduction

Multi-agent reinforcement learning (MARL) faces a critical scalability gap: while most MARL research focuses on coordinating a small number of agents over hundreds of timesteps, large-scale real-world applications demand coordination among hundreds or thousands of agents over extended horizons spanning thousands of timesteps or life-long operation. This disparity has created a bifurcated landscape where general-purpose MARL algorithms remain less tested at scale, while methods handling large-scale coordination are typically domain-specific with limited transferability.

This research addresses the fundamental question: How can we design multi-agent learning systems that simultaneously scale to large agent teams and extended temporal horizons while maintaining generalizability? It establishes that *hierarchical heterogeneous modular architectures*, integrating modules such as self-organizing neural networks, large language models, and deep reinforcement learning within unified hierarchical frameworks, provide a principled solution to both structural and temporal scalability challenges.

This work systematically investigates two complementary scaling dimensions. *Structural scaling* addresses coordination of increasing agent numbers through hierarchical decomposition and modular architectures that partition large teams into manageable subsystems. *Temporal scaling* manages long planning horizons through temporal abstraction and task decomposition, enabling agents to reason over extended sequences without exponential complexity growth. Unlike prior hierarchical approaches that employ homogeneous learning methods across hierarchy levels, the proposed frameworks leverage heterogeneous integration where each computational paradigm operates in its domain of strength: language models provide zero-shot strategic planning, self-organizing networks offer interpretable coordination, and deep reinforcement learning handles high-frequency direct interaction with environments.

This paper introduces four major contributions: a comprehensive survey establishing a novel taxonomy of MARL methods regarding scalability, a benchmark for long-horizon multi-objective multi-agent coordination (Geng et al. 2024a, 2025b), a tri-level multi-agent framework that enables deep control hierarchies through hierarchical integration of self-organizing networks with MARL (Geng et al. 2024b), and a hierarchical multi-agent framework achieving significant training sample reduction through large language model integration with MARL agents (Geng et al. 2025a). The ongoing work extends these principles to robotic tasks, investigating multi-agent approaches for coordinated manipulations.

Preliminary Research

This work addresses multi-agent scalability through four contributions that systematically progress from theoretical foundations to practical implementations. Each contribution builds upon insights from the previous work, collectively establishing hierarchical heterogeneous modular architectures as the principled solution to both structural and temporal scalability challenges (Geng 2024, 2025).

A Taxonomic Survey for Theoretical Foundations

This research conducted a comprehensive survey of 120+ papers (2015–2025) that systematically analyzes scaling approaches in MARL, establishing a novel taxonomy based

---

on control architectures and policy structures. It reveals that hierarchical heterogeneous modular architectures represent the natural evolutionary pathway for achieving scalability. The survey offers a structured perspective on multi-agent system design, identifying that successful scaling requires integrating diverse computational paradigms within unified hierarchical frameworks.

A Benchmark for Multi-Objective MARL

To address the critical gap in evaluating MARL methods on realistic scenarios, we developed the Multi-Objective StarCraft Multi-Agent Challenge (MOSMAC) (Geng et al. 2025b). MOSMAC extends planning horizons to over thousands of timesteps while incorporating multiple competing objectives, including combat effectiveness, navigation efficiency, and unit safety. Extensive benchmarking revealed counterintuitive insights: independent learning can outperform centralized methods in challenging settings, and existing methods suffer dramatic performance degradation over long horizons. These findings directly motivated the hierarchical decomposition approaches in the subsequent work.

Self-Organizing Neural Networks for MARL

Building on survey insights and MOSMAC findings, we present HiSOMA (Geng et al. 2024b) to demonstrate structural scalability through hierarchical decomposition. HiSOMA introduces a three-level architecture: the FALCON self-organizing neural network provides interpretable high-level task decomposition, middle-level *mini-MAS* modules employ state-of-the-art MARL algorithms for cooperative subtask execution, and low-level reinforcement learning policies interact directly with the environment. This heterogeneous integration, where different computational paradigms operate at their respective strengths, achieved high success rates in scenarios where flat architectures struggle, validating the hierarchical principle while revealing the need for more flexible high-level planning.

Language-Guided Zero-Shot Multi-Agent Planning

To achieve temporal scalability and planning flexibility beyond domain-specific methods, we develop the L2M2 (Geng et al. 2025a) framework. L2M2 employs large language models (LLMs) as domain-agnostic high-level planners and introduces a critical environment translator mechanism that bridges semantic planning and reinforcement learning policy execution. This design enables zero-shot generalization with high success rates while requiring significantly fewer baseline training samples, representing a significant computational reduction. This hierarchical integration of LLMs and RL demonstrates that language-guided planning can drastically reduce training sample requirements while enabling zero-shot generalization.

Future Work

Building on the presented contributions, this work continues to develop unified multi-agent architectures that integrate heterogeneous learning paradigms, where agents specialize in distinct cognitive functions, including perception, memory, reasoning, and motion, leveraging the complementary strengths through coordinated interactions. Furthermore, this research will be extended to include methods for explaining multi-agent systems, grounded in our recent studies (Wai et al. 2023, 2024).

Conclusion

This dissertation demonstrates that hierarchical heterogeneous modular architectures provide a principled solution to multi-agent scalability challenges. By systematically investigating both structural scaling (coordinating large agent populations) and temporal scaling (managing extended planning horizons), this work establishes that integrating diverse computational paradigms within hierarchical frameworks overcomes the barriers that have long constrained MARL to limited scales. The main contributions collectively bridge the gap between theoretical MARL research and real-world deployment requirements, providing both conceptual foundations and practical implementations that demonstrate significant improvements in scalability and sample efficiency.

Acknowledgments

This research was supported in part by the DSO National Laboratories, Singapore (Agreement No. DSOCL20200); the National Research Foundation, Singapore (AISG Award No. AISG2-RP-2020-019); Tsinghua University, Beijing (Grant No. R24119F0); and the SMU Presidential Doctoral Fellowships in Computing awarded to Minghong Geng.

References

Geng, M. 2024. Scaling Up Cooperative Multi-Agent Reinforcement Learning Systems. In *AAMAS 2024*, 2737–2739.

Geng, M. 2025. Hierarchical Frameworks for Scaling-Up Multi-Agent Coordination. In *AAMAS 2025*, 2932–2934.

Geng, M.; Pateria, S.; Subagdja, B.; Li, L.; Zhao, X.; and Tan, A.-H. 2025a. L2M2: A Hierarchical Framework Integrating Large Language Model and Multi-Agent Reinforcement Learning. In *IJCAI 2025*, 99–107.

Geng, M.; Pateria, S.; Subagdja, B.; and Tan, A.-H. 2024a. Benchmarking MARL on Long Horizon Sequential Multi-Objective Tasks. In *AAMAS 2024*, 2279–2281.

Geng, M.; Pateria, S.; Subagdja, B.; and Tan, A.-H. 2024b. HiSOMA: A Hierarchical Multi-Agent Model Integrating Self-Organizing Neural Networks with Multi-Agent Deep Reinforcement Learning. *Expert Syst. Appl.*, 252: 124117.

Geng, M.; Pateria, S.; Subagdja, B.; and Tan, A.-H. 2025b. MOSMAC: A Multi-Agent Reinforcement Learning Benchmark on Sequential Multi-Objective Tasks. In *AAMAS 2025*, 867–876.

Wai, K. P.; Geng, M.; Pateria, S.; Subagdja, B.; and Tan, A.-H. 2024. Explaining Sequences of Actions in Multi-Agent Deep Reinforcement Learning Models. In *AAMAS 2024*, 2537–2539.

Wai, K. P.; Geng, M.; Subagdja, B.; Pateria, S.; and Tan, A.-H. 2023. Towards Explaining Sequences of Actions in Multi-Agent Deep Reinforcement Learning Models. In *AAMAS 2023*, 2325–2327.