Hierarchical Frameworks for Scaling-up Multi-agent Coordination

Doctoral Consortium

Minghong Geng
Singapore Management University
Singapore, Singapore
mhgeng.2021@phdcs.smu.edu.sg

ABSTRACT

Multi-agent reinforcement learning has emerged as a powerful framework for developing collaborative behaviors in autonomous systems. However, existing MARL methods often struggle with scalability in terms of both the number of agents and decision-making horizons. My research focuses on developing hierarchical approaches to scale up MARL systems through two complementary directions: structural scaling by increasing the number of coordinated agents and temporal scaling by extending planning horizons. My initial work introduced HiSOMA, a hierarchical framework integrating self-organizing neural networks with MARL for long-horizon planning, and MOSMAC, a benchmark for evaluating MARL methods on multi-objective MARL scenarios. Building on these foundations, my recent work studies L2M2, a novel framework that leverages large language models for high-level planning in hierarchical multi-agent systems. My ongoing research explores complex bimanual control tasks, specifically investigating multi-agent approaches for coordinated dual-hand manipulation.

KEYWORDS

Multi-agent Reinforcement Learning; Hierarchical Multi-agent System; Large Language Model; Benchmark

ACM Reference Format:

Minghong Geng . 2025. Hierarchical Frameworks for Scaling-up Multi-agent Coordination: Doctoral Consortium. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

Multi-agent reinforcement learning (MARL) has emerged as a powerful framework for developing collaborative multi-agent systems (MAS), demonstrating remarkable success across various domains including traffic control [33], game playing [25], and robotic manipulation [22]. However, scaling up MARL for large numbers of agents and extended planning horizons remains a fundamental challenge. Current approaches face several critical limitations: the curse of dimensionality in large-scale systems [3], complex credit assignment over long time horizons [1], and the need for efficient exploration in high-dimensional state-action spaces [15].

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

Traditional MARL methods have primarily focused on scenarios with limited agents and relatively short horizons [5, 9]. While these approaches have shown promising results in controlled environments, they often struggle to address real-world applications that demand coordination among large agent populations over long horizons. Recent work suggests that hierarchical approaches offer a promising direction for scaling up MAS [2, 30], but developing effective hierarchical frameworks that can handle both structural and temporal scaling remains an open challenge.

My dissertation focuses on addressing these limitations through hierarchical MARL, focusing on two dimensions. The first dimension focuses on structural scaling, where we develop methods to effectively coordinate increasing numbers of agents through hierarchical decomposition and modular architectures, enabling MARL algorithms to handle larger agent populations while maintaining computational efficiency. The second dimension addresses temporal scaling, where we create novel approaches for managing long planning horizons and complex sequential tasks through temporal abstraction and task decomposition. By systematically investigating both dimensions, this research aims to advance MARL's capability to address real-world applications that demand sophisticated coordination among numerous agents over long horizon, such as urban traffic control [33] and multi-robot coordination [22].

2 PRELIMINARY STUDIES

My preliminary research [10] investigates scaling up MARL systems through two complementary studies: HiSOMA [12], a hierarchical control framework and MOSMAC [11], a multi-objective MARL benchmark [23]. Due to the page limit, we refer interested readers to my earlier doctoral consortium paper published at AAMAS 2024 and the full papers for these works for more details.

2.1 HiSOMA: A Hierarchical Framework for Multi-agent Hierarchical Coordination

HiSOMA [12] is a novel three-level control architecture where a FALCON controller [28] coordinates multiple subsystems known as *mini-MAS* to address the challenges of MARL in long-horizon planning [7, 15]. This design enables both structural decomposition through modular mini-MAS units and temporal decomposition through sequential subtask allocation. HiSOMA employs state-of-the-art MARL algorithms including QMIX [24], QTRAN [26], and Qatten [31] for middle-level control while using learned cognitive codes for high-level planning. HiSOMA demonstrated significant improvements over baseline approaches, particularly in complex scenarios where non-hierarchical methods struggle [20, 30].

2.2 MOSMAC: Benchmarking MARL on Long-horizon Multi-objective Tasks

Our research revealed a critical gap in the MARL community: the lack of comprehensive benchmarks for evaluating methods in long-horizon [14] and multi-objective [23] scenarios. We addressed this limitation by developing MOSMAC [11], which extends the SMAC [25] to include multiple objectives including *combat*, *navigation*, and *safety*. MOSMAC introduces scenarios where agents must balance multiple competing objectives [23] while maintaining coordination over long horizons [19]. Through extensive evaluation of state-of-the-art MARL algorithms, MOSMAC revealed critical limitations in current approaches to long-horizon planning and multi-objective optimization. These findings have helped identify promising directions for developing more scalable MARL systems.

3 FUTURE WORK

Building upon previous preliminary studies, I am pursuing several research directions to further advance the scalability and applicability of MAS. My ongoing work focuses on two key areas: investigating efficient architectures for multi-agent coordination and developing novel approaches for complex bimanual control tasks. These directions aim to advance hierarchical MARL while addressing practical challenges in real-world applications.

3.1 L2M2: Language-guided Hierarchical MARL

Recent studies in hierarchical MARL have demonstrated promising results in coordinating multiple agents [2, 18], but often rely heavily on domain-specific heuristics and pre-defined subtasks [13, 20]. Large language models have shown remarkable capabilities in reasoning and planning [8, 16], though their integration with MARL remains largely unexplored. We developed L2M2 (**L**arge **L**anguage **M**odel and **M**ulti-agent Reinforcement Learning), a novel framework that leverages LLMs for high-level strategic planning in MAS.

L2M2 introduces a hierarchical architecture where an LLM agent serves as a domain-agnostic planner, following the *LLM-as-policy* paradigm [6, 27] to decompose complex tasks and coordinate RL agents. At the framework's core is a specialized translator module that enables bidirectional communication between the LLM and MARL components, converting environmental states into natural language for LLM processing and translating the LLM's decisions into concrete subtasks. This approach builds upon recent advances in combining language models with embodied agents [32] while addressing the unique challenges of multi-agent coordination.

A key innovation of L2M2 is its *zero-shot* planning capability [8], where pre-trained LLMs can effectively guide MARL agents without requiring additional training on specific tasks. This significantly reduces the computational overhead typically associated with hierarchical MARL methods [21, 30]. We demonstrated that L2M2 achieves superior performance while requiring significantly fewer training samples compared to existing approaches [24, 26]. The framework maintains robust performance without pre-defined waypoints or heuristics as guidance, demonstrating effective strategy generation through language-based reasoning [4].

These results highlight how integrating LLMS with MARL systems creates a powerful synergy for scaling up multi-agent coordination. By combining LLMs' strategic reasoning capabilities with MARL's precise control abilities, L2M2 opens new avenues for developing more generalizable and efficient MAS [17, 34]. Our work also suggests promising directions for future research in language-guided policy learning and multi-agent coordination.

3.2 MARL for Bimanual Control

Another focus of my research is multi-agent approaches for complex bimanual control tasks, particularly on dexterous musical performance. While existing methods for hand motion synthesis have shown success in object grasping and manipulation, coordinated dual-hand control remains challenging due to its high dimensionality and complex temporal synchronization requirements. Most current approaches rely heavily on motion capture data or focus solely on single-hand control, limiting their applicability and scalability. We propose reframing bimanual control as a hierarchical multi-agent coordination problem, where different components of the hands are modeled as cooperative agents within a three-level control hierarchy: body-level coordination, hand-level management, and finger/joint-level execution.

Our initial work focuses on guitar playing as a representative test case, leveraging our experience with L2M2's language-guided planning for high-level task decomposition and HiSOMA's hierarchical control for temporal coordination. This domain presents unique challenges in terms of heterogeneous task requirements between hands and precise temporal synchronization [29]. By extending our hierarchical MARL approaches to this physically-constrained, high-dimensional domain, we aim to develop more generalizable frameworks for complex manipulation tasks while advancing our understanding of scalable multi-agent coordination.

4 CONCLUSION

This paper presents several studies on scaling up MAS through hierarchical methodologies. My completed work, including HiSOMA and MOSMAC, has established frameworks for hierarchical control and evaluation in MARL systems, while the recent development of L2M2 demonstrates how integrating LLM with MARL can create more flexible and efficient coordination. The proposed research advances the field through novel hierarchical frameworks that effectively combine different AI paradigms, scalable approaches for long-horizon coordination, and comprehensive evaluation benchmarks. Our ongoing work on bimanual control further extends these principles to complex physical manipulation tasks, validating the effectiveness of hierarchical approaches in real-world applications. As autonomous systems become increasingly prevalent, this work takes important steps toward enabling effective coordination of large-scale MAS while opening new avenues for future research.

ACKNOWLEDGMENTS

This research was supported in part by: DSO National Laboratories, Singapore (Agreement No. DSOCL20200) for the HiSOMA project and National Research Foundation, Singapore, under its AI Singapore Programme (AISG Award No: AISG2-RP-2020-019) for the MOSMAC project; conducted in collaboration with: DSO National Laboratories, Singapore for the MOSMAC project and Tsinghua University, Beijing National Research Center for Information Science and Technology for the research on L2M2 and bimanual control.

REFERENCES

[1] Adrian K. Agogino and Kagan Tumer. 2004. Unifying Temporal and Structural Credit Assignment Problems. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2 (AAMAS '04)*. IEEE Computer Society, USA, 980–987.

[2] Sanjeevan Ahilan and Peter Dayan. 2019. Feudal Multi-Agent Hierarchies for Cooperative Reinforcement Learning. *arXiv preprint arXiv:1901.08492 [cs.MA]* (Jan. 2019). https://doi.org/10.48550/arXiv.1901.08492 arXiv:1901.08492 [cs.MA]

[3] Lucian Buşoniu, Robert Babuška, and Bart De Schutter. 2008. A Comprehensive Survey of Multiagent Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38, 2 (March 2008), 156–172. https://doi.org/10.1109/TSMCC.2007.913919

[4] Yuji Cao, Huan Zhao, Yuheng Cheng, Ting Shu, Guolong Liu, Gaoqi Liang, Junhua Zhao, and Yun Li. 2024. Survey on Large Language Model-Enhanced Reinforcement Learning: Concept, Taxonomy, and Methods. https://doi.org/10.48550/arXiv.2404.00282 arXiv:2404.00282 [cs]

[5] Micah Carroll, Rohin Shah, Mark K. Ho, Thomas L. Griffiths, Sanjit A. Seshia, Pieter Abbeel, and Anca Dragan. 2019. On the Utility of Learning about Humans for Human-AI Coordination. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Curran Associates Inc., Red Hook, NY, USA, 5174–5185.

[6] Thomas Carta, Clément Romac, Thomas Wolf, Sylvain Lamprier, Olivier Sigaud, and Pierre-Yves Oudeyer. 2023. Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning. In *Proceedings of the 40th International Conference on Machine Learning (ICML'23, Vol. 202)*. JMLR.org, Honolulu, Hawaii, USA, 3676–3713.

[7] Filippos Christianos, Lukas Schäfer, and Stefano V. Albrecht. 2021. Shared Experience Actor-Critic for Multi-Agent Reinforcement Learning. https://doi.org/10.48550/arXiv.2006.07169 arXiv:2006.07169 [cs]

[8] Murtaza Dalal, Tarun Chiruvolu, Devendra Chaplot, and Ruslan Salakhutdinov. 2024. Plan-Seq-Learn: Language Model Guided RL for Solving Long Horizon Robotics Tasks. https://doi.org/10.48550/arXiv.2405.01534 arXiv:2405.01534 [cs]

[9] Benjamin Ellis, Jonathan Cook, Skander Moalla, Mikayel Samvelyan, Mingfei Sun, Anuj Mahajan, Jakob N. Foerster, and Shimon Whiteson. 2023. SMACv2: An Improved Benchmark for Cooperative Multi-Agent Reinforcement Learning. *arXiv preprint arXiv:2212.07489v2 [cs.LG]* (Oct. 2023). https://doi.org/10.48550/arXiv.2212.07489

[10] Minghong Geng. 2024. Scaling up Cooperative Multi-agent Reinforcement Learning Systems. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS '24)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2737–2739.

[11] Minghong Geng, Shubham Pateria, Budhitama Subagdja, and Ah-Hwee Tan. 2024. Benchmarking MARL on Long Horizon Sequential Multi-Objective Tasks. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS '24)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2279–2281.

[12] Minghong Geng, Shubham Pateria, Budhitama Subagdja, and Ah-Hwee Tan. 2024. HiSOMA: A Hierarchical Multi-Agent Model Integrating Self-Organizing Neural Networks with Multi-Agent Deep Reinforcement Learning. *Expert Systems with Applications* 252 (Oct. 2024), 124117. https://doi.org/10.1016/j.eswa.2024.124117

[13] Nikunj Gupta, G. Srinivasaraghavan, Swarup Mohalik, Nishant Kumar, and Matthew E. Taylor. 2023. HAMMER: Multi-level Coordination of Reinforcement Learning Agents via Learned Messaging. *Neural Computing and Applications* (Oct. 2023). https://doi.org/10.1007/s00521-023-09096-6

[14] Conor F. Hayes, Roxana Rădulescu, Eugenio Bargiacchi, Johan Källström, Matthew Macfarlane, Mathieu Reymond, Timothy Verstraeten, Luisa M. Zintgraf, Richard Dazeley, Fredrik Heintz, Enda Howley, Athirai A. Irissappane, Patrick Mannion, Ann Nowé, Gabriel Ramos, Marcello Restelli, Peter Vamplew, and Diederik M. Roijers. 2022. A Practical Guide to Multi-Objective Reinforcement Learning and Planning. *Autonomous Agents and Multi-Agent Systems* 36, 1 (April 2022), 26. https://doi.org/10.1007/s10458-022-09552-y

[15] Pablo Hernandez-Leal, Michael Kaisers, Tim Baarslag, and Enrique Munoz de Cote. 2019. A Survey of Learning in Multiagent Environments: Dealing with Non-Stationarity. *arXiv preprint arXiv:1707.09183v2 [cs.MA]* (March 2019). arXiv:1707.09183v2

[16] Sirui Hong, Mingchen Zhuge, Jonathan Chen, Xiawu Zheng, Yuheng Cheng, Jinlin Wang, Ceyao Zhang, Zili Wang, Steven Ka Shing Yau, Zijuan Lin, Liyang Zhou, Chenyu Ran, Lingfeng Xiao, Chenglin Wu, and Jürgen Schmidhuber. 2023. MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework. In *ICLR 2024*.

[17] Bin Hu, Chenyang Zhao, Pu Zhang, Zihao Zhou, Yuanhang Yang, Zenglin Xu, and Bin Liu. 2024. Enabling Intelligent Interactions between an Agent and an LLM: A Reinforcement Learning Approach. https://doi.org/10.48550/arXiv.2306.03604 arXiv:2306.03604 [cs]

[18] Shariq Iqbal, Robby Costales, and Fei Sha. 2022. ALMA: Hierarchical Learning for Composite Multi-Agent Tasks. *Advances in Neural Information Processing Systems* 35 (Dec. 2022), 7155–7166.

[19] Nan Jiang and Alekh Agarwal. 2018. Open Problem: The Dependence of Sample Complexity Lower Bounds on Planning Horizon. In *Proceedings of the 31st Conference On Learning Theory*. Proceedings of Machine Learning Research, Vol. 75. PMLR, 3395–3398.

[20] Xiangyu Kong, Bo Xin, Fangchen Liu, and Yizhou Wang. 2017. Revisiting the Master-Slave Architecture in Multi-Agent Deep Reinforcement Learning. *arXiv preprint arXiv:1712.07305v1 [cs.AI]* (Dec. 2017). https://doi.org/10.48550/arXiv.1712.07305 arXiv:1712.07305v1 [cs.AI]

[21] Saurabh Kumar, Pararth Shah, Dilek Hakkani-Tur, and Larry Heck. 2017. Federated Control with Hierarchical Multi-Agent Deep Reinforcement Learning. *arXiv preprint arXiv:1712.08266v1 [cs.AI]* (Dec. 2017). https://doi.org/10.48550/arXiv.1712.08266 arXiv:1712.08266v1 [cs.AI]

[22] Afshin Oroojlooy and Davood Hajinezhad. 2023. A Review of Cooperative Multi-Agent Deep Reinforcement Learning. *Applied Intelligence* 53, 11 (June 2023), 13677–13722. https://doi.org/10.1007/s10489-022-04105-y

[23] Roxana Rădulescu, Patrick Mannion, Diederik M. Roijers, and Ann Nowé. 2019. Multi-Objective Multi-Agent Decision Making: A Utility-Based Analysis and Survey. *Autonomous Agents and Multi-Agent Systems* 34, 1 (Dec. 2019), 10. https://doi.org/10.1007/s10458-019-09433-x

[24] Tabish Rashid, Mikayel Samvelyan, Christian Schroeder de Witt, Gregory Farquhar, Jakob Foerster, and Shimon Whiteson. 2018. QMIX: Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning. *arXiv:1803.11485 [cs, stat]* (June 2018). arXiv:1803.11485 [cs, stat]

[25] Mikayel Samvelyan, Tabish Rashid, Christian Schroeder de Witt, Gregory Farquhar, Nantas Nardelli, Tim G. J. Rudner, Chia-Man Hung, Philip H. S. Torr, Jakob Foerster, and Shimon Whiteson. 2019. The StarCraft Multi-Agent Challenge. *arXiv preprint arXiv:1902.04043v5 [cs.LG]* (Dec. 2019). arXiv:1902.04043v5 [cs.LG]

[26] Kyunghwan Son, Daewoo Kim, Wan Ju Kang, David Earl Hostallero, and Yung Yi. 2019. QTRAN: Learning to Factorize with Transformation for Cooperative Multi-Agent Reinforcement Learning. *arXiv:1905.05408 [cs, stat]* (May 2019). arXiv:1905.05408

[27] Andrew Szot, Max Schwarzer, Harsh Agrawal, Bogdan Mazoure, Walter Talbott, Katherine Metcalf, Natalie Mackraz, Devon Hjelm, and Alexander Toshev. 2024. Large Language Models as Generalizable Policies for Embodied Tasks. https://doi.org/10.48550/arXiv.2310.17722 arXiv:2310.17722 [cs]

[28] Ah-Hwee Tan, Budhitama Subagdja, Di Wang, and Lei Meng. 2019. Self-Organizing Neural Networks for Universal Learning and Multimodal Memory Encoding. *Neural Networks* 120 (Dec. 2019), 58–73. https://doi.org/10.1016/j.neunet.2019.08.020

[29] Pei Xu and Ruocheng Wang. 2024. Synchronize Dual Hands for Physics-Based Dexterous Guitar Playing. In *SIGGRAPH Asia 2024 Conference Papers (SA Conference Papers '24)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3680528.3687692

[30] Jiachen Yang, Igor Borovikov, and Hongyuan Zha. 2020. Hierarchical Cooperative Multi-Agent Reinforcement Learning with Skill Discovery. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '20)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1566–1574.

[31] Yaodong Yang, Jianye Hao, Ben Liao, Kun Shao, Guangyong Chen, Wulong Liu, and Hongyao Tang. 2020. Qatten: A General Framework for Cooperative Multiagent Reinforcement Learning. https://doi.org/10.48550/arXiv.2002.03939 arXiv:2002.03939 [cs]

[32] Bangguo Yu, Hamidreza Kasaei, and Ming Cao. 2023. Co-NavGPT: Multi-Robot Cooperative Visual Semantic Navigation Using Large Language Models. https://doi.org/10.48550/arXiv.2310.07937 arXiv:2310.07937 [cs]

[33] Huichu Zhang, Siyuan Feng, Chang Liu, Yaoyao Ding, Yichen Zhu, Zihan Zhou, Weinan Zhang, Yong Yu, Haiming Jin, and Zhenhui Li. 2019-05-13/2019-05-17. CityFlow: A Multi-Agent Reinforcement Learning Environment for Large Scale City Traffic Scenario. In *The World Wide Web Conference (WWW '19)*. Association for Computing Machinery, New York, NY, USA, 3620–3624. https://doi.org/10.1145/3308558.3314139

[34] Zihao Zhou, Bin Hu, Chenyang Zhao, Pu Zhang, and Bin Liu. 2024. Large Language Model as a Policy Teacher for Training Reinforcement Learning Agents. In *IJCAI 2024*. arXiv. https://doi.org/10.48550/arXiv.2311.13373 arXiv:2311.13373 [cs]