Scaling up Cooperative Multi-agent Reinforcement Learning Systems

Minghong Geng

Neural and Cognitive Computing Group
Singapore Management University
Email: mhgeng.2021@phdcs.smu.edu.sg
Homepage: https://gengminghong.github.io

My Homepage

Research Group

Introduction

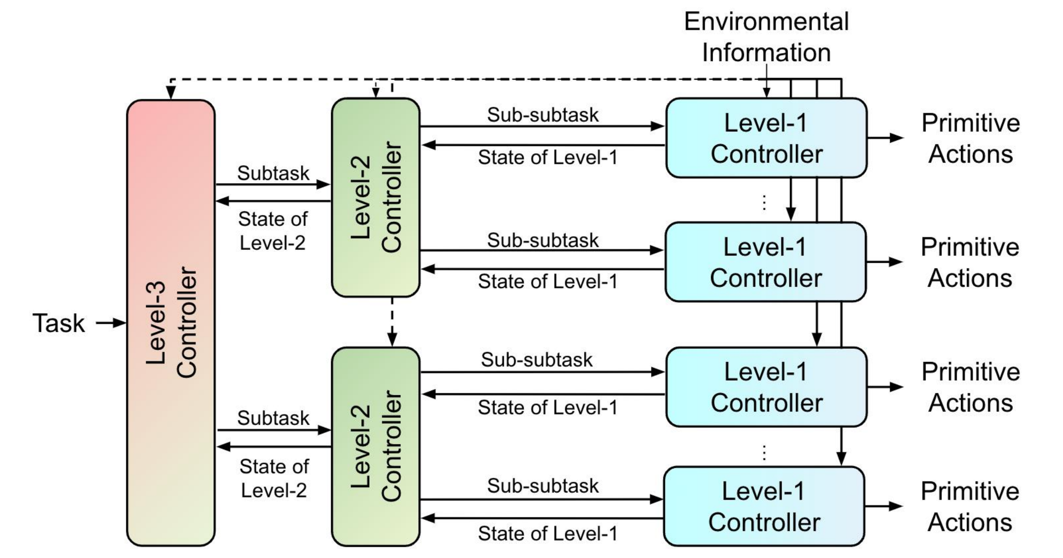
- Cooperative multi-agent reinforcement learning (MARL) methods aim to learn effective collaborative behaviours of multiple RL agents.
- While many methods have achieved notable success, a critical focus within the community is the ***scalability*** of MARL systems.
  - Challenges: *structural and temporal credit assignment, non-stationarity,* the *curse of dimensionality, lack of benchmarks, etc.*
- My studies focus on the scalability of MARL in two directions through *task decomposition:*

Two directions of scaling up MARL	***Two Task Decomposition Approaches***	***Preliminary Work***
Structural-wise Increasing the number of agents in multi-agent systems (MAS) may boost their ability to tackle complex problems that are unmanageable by smaller-scale systems.	*Structural task decomposition* To divide large-scale MAS into smaller autonomous modules called **mini-MAS** and develop effective cooperation mechanisms among mini-MAS.	**HiSOMA** A Hierarchical MARL Framework
Temporal-wise Many real-world problems involve thousands or infinite decision-making steps to complete. Increasing the number of decision-making steps for long-horizon planning is crucial.	*Temporal task decomposition* To divide long-horizon, complex problems into **subproblems** with shorter horizons and ensure agents learn generalizable policies across subproblems.	**MOSMAC** A Multi-objective MARL Benchmark

HiSOMA Overview

- HiSOMA is a hybrid hierarchical MARL framework that combines
  - a class of Self-Organizing Neural Network (SONN), named Fusion Architecture for Learning, Cognition, and Navigation (FALCON) [1] and
  - state-of-the-art non-hierarchical MARL methods, such as QMIX, VDN, QTRAN and Qatten

  to navigate long-horizon decision-making problems.
- HiSOMA is a novel attempt to incorporate **heterogenous learning algorithms** into a unified framework. It is verified to be able to generalized to a broad-range of MARL methods.



An illustration of a three-level HiSOMA framework. The higher-level controllers aggregate the local states from each subordinate lower-level controller and allocate decomposed subtasks. The top-level (Level-3) controller commands multiple lower-level multi-agent systems (MAS), which could be recognized as "mini-MAS" within a scaled-up MAS.

Please cite HiSOMA by: Geng, M., Pateria, S., Subagdja, B., & Tan, A. H. (2024). HiSOMA: A hierarchical multi-agent model integrating Self-Organizing Neural Networks with multi-agent deep reinforcement learning. ***Expert Systems with Applications***, 124117.

MOSMAC Overview (More details will be shown at Poster stall 1B on 9th May 2024)

❖ MOSMAC features three characteristics:

1 **M**ultiple **O**bjectives

Objective 1 (combat): The damage to enemy units should be as much as possible.

Objective 2 (navigate): The distance agents from agents to strategic positions should be as minimal as possible.

2 **S**equential **T**ask **A**llocation

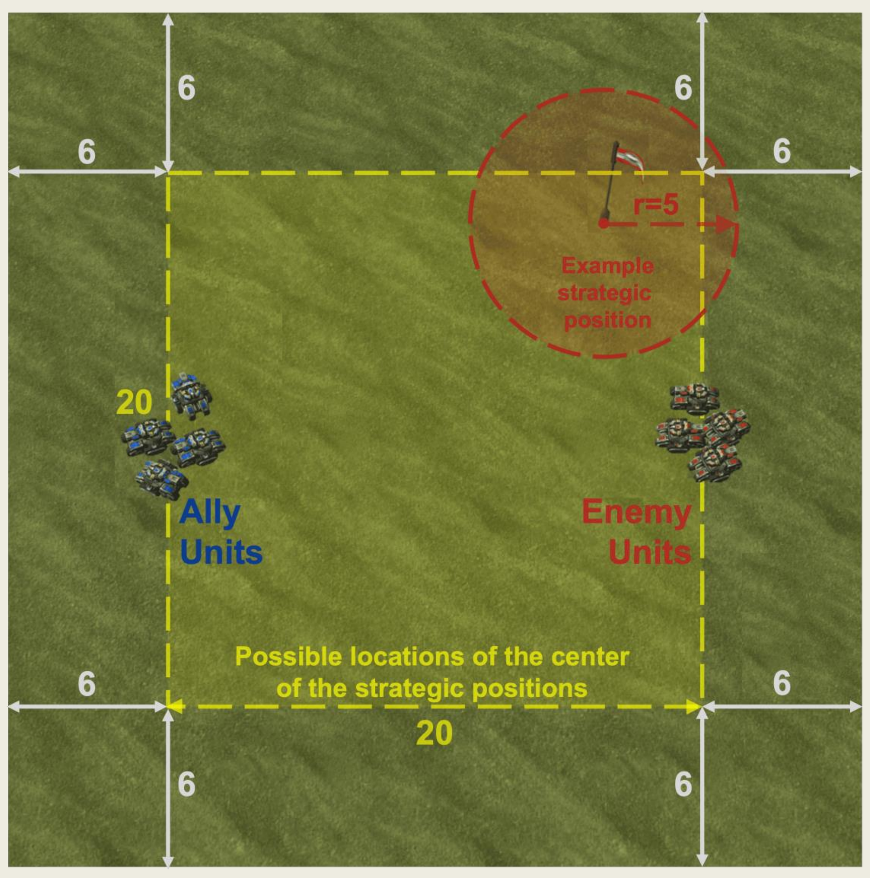
Agents are allocated with tasks sequentially in an episode, where the completion of the previous task triggers the next task allocation. As such, the long horizon tasks are decomposed into many subtasks.

3 **V**arying **H**orizons

The total timesteps of each long-horizon task vary based on the target locations and paths. The environment generates random directed acyclic paths for agents to navigate from the starting to the target locations.

❖ MOSMAC contains two sets of scenarios:

Short-horizon MOSMAC



An illustration of a short-horizon MOSMAC scenario named 4t. The winning condition is all alive agents arrive at the strategic position, motivated by two objectives. The target locations are randomly selected for each episode, inspired by SMACv2 [2].

Long-horizon MOSMAC

An illustration of a long-horizon MOSMAC scenario, named 4t_vs_12t. The length of the horizons is configurable by moving the final targets to any of the red areas. Enemies are placed in three positions to deter the advancement of multi-objective agents.

Please cite MOSMAC by: 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.

Findings in Preliminary Results

- We evaluate nine MARL methods on the MOSMAC benchmark with the EPyMARL [3] framework.
- Existing MARL methods can address short-horizon tasks but struggle when dealing with sequential tasks that involve multiple objectives over a longer horizon.
- Long-horizon multi-objective learning poses significant challenges for MARL algorithms, as evident from the performance drops in long-horizon scenarios compared to short-horizon ones.
- The results of HiSOMA further illustrate the potential of applying hierarchical MARL methods for addressing long-horizon problems, which are hard to address by MARL methods without applying temporal and structural task decomposition.

Towards Large Agent Team and Long-Horizon Planning

- Moving forward, my dissertation aims to broaden the scope by exploring two pivotal research directions of MARL.
- Future work includes expanding MOSMAC with additional objectives, scenarios, and algorithms, as well as exploring hierarchical learning and domain knowledge-based task decomposition to improve performance on long-horizon multi-objective tasks.
- It is also important to develop benchmarks and tasks tailored for scaled-up MARL methods and to establish standard evaluation criteria.

References

[1] 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.

[2] B. Ellis et al., "SMACv2: An Improved Benchmark for Cooperative Multi-Agent Reinforcement Learning", arXiv preprint arXiv:2212.07489v2 [cs.LG].

[3] G. Papoudakis, F. Christianos, L. Schäfer, and S. V. Albrecht, "Benchmarking Multi-Agent Deep Reinforcement Learning Algorithms in Cooperative Tasks", in Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks (2020).