Project Proposal : Undergraduate Research

Hierarchical Management of Multi-Agent Teams
Timeline: UGR 4971/4972 to be completed Spring 2025 and Fall 2025.

Hugo Garrido-Lestache, CS

March 13, 2025

1   Project Goals

Reinforcement learning holds great promise for training AI agents capable of managing complicated control tasks with long term temporal dependence between the actions of the agent and the evolution of the environment. Many real world tasks involve high-dimensional state-dependent action spaces and/or sparse reward structures, and it can be challenging to manage these tasks effectively using agents trained with traditional classical or modern reinforcement learning algorithms.

Hierarchical reinforcement learning (HRL) is a promising framework that enables agents to tackle complex, multi-step tasks by breaking them down into manageable subtasks. This decomposition into hierarchies is especially valuable in environments with sparse or delayed rewards, where traditional methods struggle. Given this promise, ongoing research in HRL is focused on at least three areas. First, currently, most HRL algorithms are bottom-up. The agent first learns sub-behaviors, and then once these sub-behaviors are well developed, learns how to compose them to solve complex problems. The bottom-up approach can be inefficient if the agent spends time learning sub-behaviors it does not need to solve the task at hand. An alternative is the top-down approach, which decomposes the problem first and is more sample efficient, but leads to unstable learning [1, 4, 5] while a variety of non-stationary, model-based, or policy-based approaches have been proposed to deal with this [4,3,7]. Second is the issue of goal representation. Approaches to this range from use of the full state-space [3] which is difficult to scale, to more compact representations [4] learned end-to-end which reduces the amount of different sub-behaviors which can be expressed, and finally to expressing it as an optimization problem [8]. An optimal goal representation will likely be lower-dimensional than the full state-space, and use some degree of abstraction. Methods to learn the right abstraction only using data gathered from the environment are unknown. Finally, most HRL approaches [1,2,3,4], make use of a two-layered approach in which a manager chooses sub-behaviors to activate where those sub-behaviors decide on primary actions. It is speculated [5,6,7] that algorithms in HRL may benefit from using more than two levels, especially for complex long term planning.

In this project I will study the use of hierarchical reinforcement learning techniques to manage collaborative team tasks. Specifically, this research will aim to add to our understanding of collaborative multi-agent systems by directly comparing the performance of an unmanaged group of trained agents vs a suite of managed groups constructed using various forms of HRL drawn from bottom-up vs top-down approaches,

differing granularities of goal-representation, and deeper layers of hierarchy. I anticipate using a simple benchmark environment inspired by team sports, in this case soccer, a game in which two teams of agents play against each other by attempting to move a single ball into two set of opposing goals. This problem involves rewards (e.g. winning, scoring goals) that are sparse relative to the frequency of actions, subtasks whose successful completion are only weakly correlated with those rewards, and complex challenges including coordination between agents who are themselves endowed with varied actions.

2 Timeline

- Week 1 - 4: Literature review and survey of extant HRL techniques;

- Week 5 - 8: Building environment and Implementing traditional modern RL algorithms;

- Week 9 - 12: Implementing HRL algorithms and conducting initial experiments;

- Week 13 - 14: Identifying candidate strategies for algorithm improvement;

- Week 15 - 16: Implementing competitive league play and conducting experiments;

- Week 16 - 20: Compare and contrast model with existing approaches;

- Week 21 - 30: Prepare written thesis document, research poster, public presentation at MSOE.

3 References

1. Bacon, P.L.; Harb, J.; Precup, D. "The option-critic architecture." In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, 4–9 February 2017.

2. Eysenbach, B.; Gupta, A.; Ibarz, J.; Levine, S. "Diversity Is All You Need: Learning Skills without a Reward Function." *arXiv* 2019, arXiv:1802.06070.

3. Nachum, O.; Gu, S.S.; Lee, H.; Levine, S. "Data-efficient hierarchical reinforcement learning." In *Proceedings of the Advances in Neural Information Processing Systems*, Montreal, QC, Canada, 3–8 December 2018.

4. Vezhnevets, A.S.; Osindero, S.; Schaul, T.; Heess, N.; Jaderberg, M.; Silver, D.; Kavukcuoglu, K. "Feudal networks for hierarchical reinforcement learning." *In Proceedings of the International Conference on Machine Learning*, Sydney, Australia, 6–11 August 2017.

5. Haarnoja, T.; Hartikainen, K.; Abbeel, P.; Levine, S. "Latent Space Policies for Hierarchical Reinforcement Learning." In *Proceedings of the International Conference on Machine Learning*, Stockholm, Sweden, 10–15 July 2018.

6. Fox, R.; Krishnan, S.; Stoica, I.; Goldberg, K. "Multi-level discovery of deep options." *arXiv* 2017, arXiv:1703.08294.

7. Levy, A.; Platt, R.; Saenko, K. "Hierarchical Reinforcement Learning with Hindsight." In *Proceedings of the International Conference on Learning Representations*, New Orleans, LA, USA, 6–9 May 2019.

8. Nachum, O.; Gu, S.; Lee, H.; Levine, S. "Near-Optimal Representation Learning for Hierarchical Reinforcement Learning." In *Proceedings of the Advances in Neural Information Processing Systems*, Montreal, QC, Canada, 3–8 December 2018.