Emergent Multiagent Interactions Through Quality Diversity Optimization

Senior Project, Winter and Spring 2024

Author: Siddarth Viswanathan

Advisor: Franz J. Kurfess

Abstract

In multiagent settings, it is often difficult to achieve optimal performance and maximize system reward due to the multitude of synergistic agent-agent interactions that must first occur. Exploring the entirety of a state space in a single agent environment is usually infeasible itself, and this problem is only further exacerbated in multiagent environments where the *joint-state* spaces must be explored. Quality Diversity (QD) methods like MAP-Elites offer partial solutions to this problem by learning a diverse repertoire of policies based on specified behavior space. Similarly, the Asymmetric Island Model uses a QD-like method to learn a diverse set of policies, but only learns diversity in respect to accomplishing an agent-specific task. In this work, we extend the idea of Quality Diversity to incentivize agent-agent interactions to more quickly converge to a set of strong policies.

Outline

- Background
    - Cooperative Multiagent systems are settings where multiple agents must collaborate to accomplish a task.
    - The task is defined by a system reward function, G
    - One of the main challenges in this area is how to get agents to quickly learn how to work with each other effectively
    - Cooperative Coevolutionary Algorithms are a non-gradient based, evolutionary method to accomplish the system objective
        - A policy is randomly sampled for each agent and after being put onto a team with other agents, a score is given to each agent based on G (if reward shaping methods are used, this score could differ for each agent)
    - Model-free Reinforcement Learning focuses on learning the optimal policy without access to a model of the world (experience and exploration is how agents learn)
    - In single-agent continuous state space problems, exploration becomes even more of an issue.
        - Coupled with multiple agents makes it imperative to explore the *joint-state* space which is exponentially larger

    - If agents are quickly able to learn synergistic behaviors and which classes/agents are most conducive to accomplishing the system objective, this problem can be alleviated to a significant extent.
- Related Work
    - Quality Diversity Optimization like MAP-Elites are a way to learn a diverse set of policies for problems where one optimal solution might not be enough. Instead, a set of solutions are more valuable. QD methods can be used to promote diversity in agent-agent interactions by specifying a task between different agents. The main issue with this naive approach is it requires domain knowledge about how the agents should interact together without an emergence in new strategy. (In a soccer environment, specifying that the system objective is to learn to pass a ball removes the emergence of this idea)
    - A QD-like approach is known as the Island Model which breaks down a problem into islands for agents to individually learn transferable, agent-specific skills which are then applied to the "mainland" which is the actual system objective. All of the QD optimization occurs on the islands where each agent learns a population of ways to accomplish the island objective based on the specified behavior space axes (speed of a car, number of arms on a robot, observation radius, etc.) Though an island can be constructed to promote multiagent interaction, once again this requires a prescribed task for how agents are expected to interact with each other, requiring more domain knowledge.
    - Curiosity-driven intrinsic motivation is an interesting way of exploring an environment especially in sparse-reward settings. Most work surrounding this is in single-agent settings. Further, even in multiagent settings, current methods do not formulate a domain-agnostic method of promoting agent-agent interaction.
- Features, Requirements and Evaluation Criteria
    - The main contribution of this work will be a novel framework promoting agent interactions in a domain-agnostic way
        - By applying this domain-agnostic framework, performance should be better than a set of baselines
    - User interaction can be done in twofold: through the constructed codebase or as a theoretical framework/idea
    - Performance graphs will be constructed to assess how well the proposed method works compared to current methods
- System Design + Experiments
    - Dummy environment: simple soccer domain where Agent A and B must move a ball from one point to another. If an agent holds a ball itself for longer than x seconds, the reward deteriorates.
        - This environment should be simple enough to debug, but should be fruitful enough to observe an array of emergent strategies due to interaction between A and B
    - Peter Stone's Robot Soccer/Google Football would be a good environment to test this method due to the richness and importance of multiagent interactions

    - Variations of Kagan Tumer's Rover domain could be tested where agent-agent actions are required to receive reward.
- Future Work
    - Depending on the nature of the framework, the ideas could be transferred to the Island Model as another level between the Islands and Mainland.
    - Natasha Jaques introduced an idea of social influence as intrinsic motivation. This notion of influence could be combined with the proposed framework to quickly find which agent classes should interact with each other

References

https://arxiv.org/pdf/1705.05363.pdf "Curiosity-driven Exploration by Self-supervised Prediction"

https://arxiv.org/pdf/1810.08647.pdf "Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning"

https://arxiv.org/pdf/2311.01829.pdf "Mix-ME: Quality-Diversity for Multi-Agent Learning"

https://dl.acm.org/doi/abs/10.1145/3583131.3590524 "Learning Synergies for Multi-Objective Optimization in Asymmetric Multiagent Systems"

https://arxiv.org/pdf/2302.10825.pdf "Curiosity-driven Exploration in Sparse-reward Multi-agent Reinforcement Learning"

https://arxiv.org/pdf/1812.07019.pdf "Malthusian Reinforcement Learning"

https://arxiv.org/pdf/2007.03102.pdf "Natural Emergence of Heterogeneous Strategies in Artificially Intelligent Competitive Teams"

https://www.sciencedirect.com/science/article/abs/pii/S0925231221015265?fr=RR-2&ref=pdf_download&rr=846d392728469688 "Diversity-augmented intrinsic motivation for deep reinforcement learning"

https://dl.acm.org/doi/pdf/10.1145/3520304.3529062 "Balancing Teams with Quality-Diversity for Heterogeneous Multiagent Coordination"

https://openreview.net/pdf?id=BkG8sjR5Km "Emergent Coordination through Competition"

https://alaworkshop2023.github.io/papers/ALA2023_paper_16.pdf "Joint Intrinsic Motivation for Coordinated Exploration in Multi-Agent Deep Reinforcement Learning"