May 2021

Identification of Emergent Collaborative Behaviors in Multi-Agent Systems

Bryson Howell
bhowel13@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/eureca

Part of the Robotics Commons

Recommended Citation

This Article is brought to you for free and open access by the Supervised Undergraduate Student Research and Creative Work at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in EURēCA: Exhibition of Undergraduate Research and Creative Achievement by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
Background
Multi-Agent Reinforcement Learning (MARL) has been used to allow groups of autonomous agents to perform complex cooperative tasks such as stochastic games.

Prior Work
• MARL-trained teams display a variety of behaviors even when trained under identical conditions [1].
• These behaviors have been observed to be significantly coordinated [2].

Hypotheses
• Coordination is a useful measure of collaboration.
• MARL methods can produce emergent collaborative strategies.

Motivation
• Recent work has used neural networks to identify dynamical systems [3].
• No method is capable of completely identifying and describing collaboration.
• Developing a process for modeling collaborative strategies can lead to AI agents that are more adaptive to new teammates and changing environments.

Research Goals
1. Explore a method for describing the group behavior of a heterogeneous team performing a predator-prey pursuit task.
   2. Assess if this method can provide insight into the collaborative strategies learned by MARL agents and inform future work.

Methods
Environment
• Agents are simulated within a 2D particle environment utilizing the OpenAI Gym library [4].
• Agents play a continuous pursuit-evasion game (Figure 2).

MARL Algorithm
• Agents are trained with the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) method [4].
• Uses the Centralized Training and Decentralized Execution (CTDE) paradigm.

Reward structure:
• Predators are equally rewarded when one of them collides with the prey.
• Prey receives the opposite reward.

Fixed Strategies
• Defined by simple algorithms
• Not coordinated with other predators
• Behaviors are easily described:
  • **Interceptor:** Agent calculates an interception trajectory based on the prey’s current velocity.
  • **Chaser:** Agent heads directly towards the prey.
• Used for comparisons with unknown RL behaviors.

Experiment Process
1. Train RL agents for 1000 episodes
2. Swap RL predators for fixed predators
3. Agents perform task for 1000 episodes
4. Collect raw performance data
5. Transform data to fit interpretable metrics

Data Exploration
• Selected several human-interpretable metrics that could describe the strategy of a predator agent.
• Applied principal component analysis (Figure 4) to determine which features were useful.
• **Key finding:** “Average Distance from Prey” feature was consistently the most variable between agents.
• Theorized that this feature is sufficient to differentiate between team strategies.

Results
• The extent of MARL agents’ ability to adapt to their teammates can be inferred by observing the probability distributions formed by teams that contain a single MARL agent.
• By comparing these distributions to those from fixed strategy teams, the behavior of each RL agent becomes human-interpretable.

Discussion
• Results indicate that MARL agents can assume different strategies to adapt to teammates.
• Investigating the process by which MARL agents are adapting to teammates can lead to an improved model of collaboration.
• This can lead to MARL agents that are able to intelligently form teams with other AI agents and human partners.

Future Work
• Train a classifier to identify team strategies by the multimodal probability distributions visible in Figure 5.
• Use findings from fixed strategy teams to investigate teams composed only of MARL agents.
• Create an agent that swaps between fixed strategies.

Works Cited

Figure 1: MARL Overview
Figure 2: Pursuit-Evasion Task
Figure 3: Experiment Flowchart
Figure 4: Principal Component Analysis Applied to Fixed-Strategy Team
Figure 5: Probability Distributions of ‘Distance from Prey’ Feature for Various Predator Teams