Project Proposal: Emergent Adversarial Behaviors via Self-Play in Multi-Agent Reinforcement Learning

Team members name	Email id
Charan Reddy Nandyala	cnand002@ucr.edu
Dhanush Chalicheemala	dchal007@ucr.edu
Satyadev Gangineni	sgang024@ucr.edu

1. Introduction and Motivation

Reinforcement Learning (RL) has shown impressive results in domains such as robotics, game playing, and autonomous systems. However, many real-world problems involve multiple agents interacting with and adapting to each other, rather than a single agent learning in isolation. This leads to the exciting area of Multi-Agent Reinforcement Learning (MARL), where agents must learn to cooperate, compete, or coexist in shared environments.

In competitive multi-agent settings, self-play where agents learn by repeatedly training against copies or evolving versions of themselves has proven to be a powerful training paradigm. It enables the emergence of increasingly sophisticated behaviors without the need for manually designed opponents or reward structures. Classic examples include AlphaGo and competitive environments in OpenAI Gym and PettingZoo.

This project proposes to explore the dynamics of self-play in an adversarial predator–prey environment, where multiple predator and prey agents learn competing strategies through interaction and adaptation. The goal is not to achieve high performance in a specific task but to understand and visualize how adversarial learning and self-play can drive the evolution of intelligent, emergent behaviors in simple simulated ecosystems.

2. Project Overview

The proposed project will develop a grid-based predator–prey environment where two types of agents, predators and prey interact with opposing objectives: predators aim to capture prey efficiently, while prey aim to survive for as long as possible.

Both agent types will be trained using reinforcement learning techniques in a self-play setting, allowing strategies to evolve dynamically as each side adapts to the other. The environment will support both cooperative and competitive learning. Rather than focusing on algorithmic complexity, the project will emphasize conceptual understanding, qualitative observation of emergent behaviors, and empirical analysis of training progression.

3. Objectives

1. Design and implement a multi-agent predator-prey environment that supports both cooperative and adversarial dynamics.
2. Apply reinforcement learning with self-play to allow predators and prey to learn adaptive strategies over time.
3. Analyze emergent behaviors and learning stability, comparing different training setups (e.g., independent vs. centralized learning).
4. Demonstrate the potential of self-play in driving open-ended skill development in competitive environments.

4. Methodology

1. Environment Creation: Develop a custom grid-world simulation with multiple agents and discrete actions (move up, down, left, right, or stay).
2. Multi-Agent Training: Use reinforcement learning algorithms such as PPO or actor-critic variants. Predators and prey will be trained alternately or concurrently using a self-play mechanism.
3. Behavioral Analysis: Evaluate learning dynamics by tracking success rates, average survival times, and emergent patterns such as cooperative hunting or evasive maneuvers.
4. Extensions (Optional / if time persists): Incorporate curriculum learning by gradually increasing task difficulty or explore simple hierarchical extensions if time permits.

5. Expected Outcomes

- Observation of emergent behaviors such as coordinated predator movement and evasive prey patterns.
- Insights into the effectiveness of self-play in producing adaptive and robust strategies.
- Qualitative and quantitative analysis of learning dynamics in adversarial MARL settings.

- A demonstration environment extendable for future research in cooperative-competitive RL.

6. Tools and Resources

- Programming Language: Python
- Libraries: PyTorch, PettingZoo, RLlib or Stable-Baselines3, NumPy, Matplotlib
- Compute Resources: Standard GPU/CPU setup (local or Colab)
- Deliverables: Source code, trained models, visualizations, and a concise project report

8. Conclusion

This project aims to explore how competitive interactions in a self-play setting can lead to complex, adaptive behaviors in multi-agent reinforcement learning. Through the predator-prey simulation, the study will highlight how simple reward functions and interaction rules can give rise to emergent intelligence and strategy formation reflecting fundamental principles of learning, adaptation, and co-evolution in artificial agents.