# Proposal: Multi-Agent Distributed Reinforcement Learning Framework

Haoming Ma (hm3070) EECS 6892

## 1. Project Summary

This project aims to develop a distributed multi-agent reinforcement learning (MARL) training framework based on a client-server architecture. In this framework, multiple clients independently train different agent policies, while the sever is responsible for collecting, distributing and evaluating policies. By leveraging parallel training and a centralized strategy pool, this approach aims to accelerate training efficiency and increase the diversity in multi-agent settings, to address the challenges of self-play stagnation in traditional training methods.

This framework will support both competitive and cooperative MARL scenarios, allowing agents to either compete against or collaborate with each other. Additionally, it accommodates both symmetric and asymmetric agent roles, enabling the study of scenarios where agents have identical capabilities as well as those where they assume distinct responsibilities. The system is designed to support scalable distributed training, allowing agents to be trained in parallel across multiple computational nodes.

## 2. Key References

L. Espeholt *et al.*, IMPALA: Scalable distributed deep-RL with importance weighted actor-learner architectures

J. Foerster *et al.* , Counterfactual Multi-Agent Policy Gradients

C. Berner *et al.* , Dota 2 with Large Scale Deep Reinforcement Learning

## 3. Key Directions & Expected Outcomes

Key Directions

- Implement a distributed training system where clients train policies while the server handles policy distribution and evaluation.
- Measure the impact of distributed training on convergence speed and policy robustness.
- Investigate how different evaluation strategies affect policy evolution in cooperative and competitive settings

Expected Outcomes

- A working prototype of a distributed MARL framework, validated on Gym environments.
- Insights into policy evaluation mechanisms and their impact on MARL training.
- Potential extension into real-world applications such as robotics or economic simulations.