COMP0036 – Project Plan:

Multiagent Reinforcement Learning for Noisy Communication in MPEs

**Jeffrey Li**, supervised by **Professor Mirco Musole**

BSc Computer Science, UCL

November 17, 2022

# Project Overview

**Aims:**

To devise an algorithm for effective communication of agents with noisy communication channels to achieve fully cooperative coordination tasks in selected Multi-Particle Environments. The algorithm would make use of Multiagent Reinforcement Learning concepts with centralized learning and decentralized execution which takes into account the following assumptions and constraints:

**Assumption and constraints**

* Tasks to solve are fully corporative and individual agents have partial observability of the environment.
* Agents can communicate with one and the other using a discrete communication channel with no limits in bandwidth.
* Learned policies can only make use information local to each agent at execution time
* The algorithm does not assume a differentiable communication channel between agents.
* The noise added to communication channel is unknown to the agents.

**Objectives:**

* 1. **Review on Reinforcement Learning, Deep Reinforcement Learning concepts.**
  2. **Research and understand underpinnings of MARL in the context of coordination through communication (with noise).**
  3. **Conduct in-depth literature review on related topics and methods**
  4. **Devise my algorithmic solution mathematically and in pseudocode.**
  5. Implement the proposed pseudocode in Python.
  6. **Train and test the Python implementation on the listed MPE environments.**
  7. **Evaluate success of implemented algorithms with other state-of-the-art MARL algorithms.**

# Expected deliverables

The expected deliverable would start with an overview of the problem to be solved, explaining its motivation and context.

I would then move onto a broad survey summarizing core concepts of Reinforcement Learning, Deep Reinforcement Learning and Multiagent Reinforcement Learning for the more general readers. The literature survey would be backed by Python 3.x implementations for the following RL and Deep RL algorithms that are trained and tested in OpenAI Gym’s toy environments:

* **Value Iteration**
* **Policy Iteration**
* **Q Learning**
* **Deep Q Learning**

The project report would then have a greater focus on cooperative MARL in the context of coordination and communication with broad literature review over the array of algorithmic solutions on topics related to my project.

I would then propose my algorithm for solving the proposed problem. This would include mathematical formulations as well as pseudocode and would also be backed by my Python 3.x implementation of the algorithm.

This implementation would be trained and tested on a selective of benchmark MPEs [1] as shown below with my modification in adding additional noises to agent’s communication channels:

* Simple Spread
* Simple Speak Listener
* Simple Reference

And is then followed by an in-depth evaluation on its performance presented in the form of figures and graphs. This would be done by comparing results against the state-of-the-art algorithms listed below which would also be trained on the same MPEs:

* TODO

The metrics used for comparison would be the mean episodic reward in training and in testing. The results would be obtained on varied number of cooperative agents with optimal (to my best ability) network hyper-parameters for each environment.

In conclusion, I will perform analysis on strengths and weaknesses of my proposed algorithm based on the evaluation and offer future work to be done as well as areas for improvements.

# Work plan

* Project start to mid-Nov (4 weeks)
  + Complete reviewing, implementing and testing on Reinforcement and Deep Reinforcement Learning algorithms
* Mid-Nov to mid-Dec (4 weeks)
  + Complete Project plan and Literature review on proposed topic
  + Modify selected MPEs for additive noises in communication
* Dec 19th to Early-Jan – Christmas Break

(3 weeks)

* Early-Jan to 18th Jan (2 weeks)
  + Work on completing the interim report
* **18th Jan - Interim Report Due**
* 18th Jan – Mid Mar (7 weeks)
  + Finalise mathematical formulation and pseudocode for my algorithm
  + Finish implementation of my algorithm.
  + Obtain training and testing results of my algorithm in the modified MPEs.
* Mid-Mar to Late-Mar (2 weeks)
  + Implement listed state-of-the-art algorithms.
  + Obtain training and testing results for these algorithms in the modified MPEs.
* Late-Mar to 26th April (5 weeks)
  + Work on completing Final Project report
* **26th April - Project Submission**

# References:

[1] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, “Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments.” arXiv, 2017. doi: 10.48550/ARXIV.1706.02275.

[2]

[3]