Multi-Agent System for Cooperative Navigation

Magdalena Warylak
111531
Instituto Superior Técnico

Jakub Krupiński
111530
Instituto Superior Técnico

José Palha
87051
Instituto Superior Técnico

ABSTRACT

Project presents a multiagent system for cooperative navigation, addressing coordination challenges and communication dynamics among autonomous agents. Leveraging the Simple Spread framework, the system orchestrates interactions to optimize task performance while minimizing collisions. Through simulations with agents and landmarks, agents learn individual policies and coordination strategies to cover landmarks efficiently. Evaluation metrics assess both local collision avoidance and global task performance.

1 Introduction

In recent years, multiagent systems have garnered significant attention due to their ability to model and solve complex problems by coordinating the actions of multiple autonomous agents. This project aims at developing a multiagent system for cooperative navigation, with the objective of addressing challenges related to coordination problems, communication and cooperation between agents.

The motivation behind this project stems from the growing need for efficient coordination and collaboration among autonomous entities in various domains. With the rise of artificial intelligence and autonomous systems, there is a pressing demand for multiagent system solutions that can adapt to dynamic environments and handle uncertainty. In many real-world scenarios, such as robotic swarm deployment, autonomous vehicle routing, and distributed sensor networks, effective coordination among agents is essential for achieving optimal outcomes.

The problem we aim to address revolves around coordinating a group of autonomous agents to perform collaborative tasks efficiently. For this purpose, we use the Spread environment from the PettingZoo framework. Specifically, the Spread environment addresses the problem of multi-agent, which need to learn both individual policies and coordination strategies to achieve optimal performance. Agents interact with each other and the environment, making decisions that impact on all agents' rewards.

Our main objective is to simulate the problem using a multiagent system based on Spread environment and test it in terms of policy given to the agents. We would like to study the behaviour of agents, based on the policies given to them and inferred using reinforcement learning.

2 Related Work

In the realm of Multi-Agent Reinforcement Learning (MARL), addressing cooperative problems has garnered considerable attention. One notable contribution is the paper titled "On Solving Cooperative MARL Problems with a Few Good Experiences." as shown in [1]. This work explores the challenges of cooperative MARL and proposes a solution based on experience replay, where agents store and reuse successful experiences to improve learning efficiency and coordination.

Another relevant study in cooperative MARL is [2]. This comprehensive review discusses various theoretical frameworks and algorithms for cooperative MARL, shedding light on the complexities and potential solutions in this domain.

3 Approach

To simulate the problem, we use an environment that has N independent agents and N landmarks. Agents must learn to cover all the static landmarks while avoiding collisions. More specifically, all agents are globally rewarded based on how far the closest agent is to each landmark (sum of the minimum distances). Locally, the agents are penalized if they collide with other agents (-1 for each collision).

> Agent observations: [self_velocity, self_position, landmark_relative_positions, other_agent_relative_positions, communication]
> Agent possible actions: [no_action, move_left, move_right, move_down, move_up]

Each agent to coordinate needs to know information about other agents' positions (to avoid collisions) as well as positions of landmarks (to know which area to cover). With that information agents can learn a policy or try one already defined to maximize the defined rewards.

Each agent can have one of three policies:

- Go straight to the nearest landmark
- Behavior-tree based policy
- Reinforcement learning based policy

4 Empirical evaluation

Defined metrics for evaluating agents' behaviour will be local and global rewards. Local rewards will indicate how often an agent collides with other agents. Global rewards show how well the task is performed by all the agents. Ideally the agents will maximize the rewards while not being penalized for collisions.

REFERENCES

[1] Rajiv Ranjan Kumar and Pradeep Varakantham, 2020. Solving Cooperative MARL Problems with a Few Good Experiences. DOI: https://doi.org/10.48550/arXiv.2001.07993.

[2] Kaiqing Zhang, Zhuoran Yang and Tamer Başar. 2021. Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms. DOI: https://doi.org/10.48550/arXiv.1911.10635.