INFO-H512 - Current trends in artificial intelligence

Swarm-Warriors: A Deep Reinforcement Learning Framework for Team-Based Agent Combat Simulation

Author :

BONHEURE Luka

MAVROS Mickail

Teacher :

BERSINI Hugues

2024

Contents

Abstract

This project presents Swarm-Warriors, a novel framework implementing deep reinforcement learning to simulate coordinated combat between two teams of agents. Each team comprises ranged and melee units whose behaviours are governed by neural networks and Q-learning algorithms. The agents make decisions based on local information, including their surroundings and health points, ensuring map-independence and localised behaviour akin to swarm intelligence. The framework facilitates easy map generation and agent characteristic modification, promoting flexibility and extensibility in combat simulations. Additionally, an application was developed to streamline testing and visualisation of simulations. All the code of the project can be found on Github.

1 Introduction

In recent years, deep reinforcement learning (DRL) has shown significant promise in complex decision-making scenarios, from games to autonomous driving. This project explores its application in simulating team-based combat scenarios, where agents must coordinate to achieve a common goal: defeating the opposing team. The unique challenge lies in the coordination of heterogeneous agents (ranged and melee) using only local information, akin to principles observed in swarm intelligence.

Swarm intelligence, observed in nature among insects like ants and bees, involves decentralised, self-organised systems where individuals follow simple rules to achieve complex group behaviours. In robotics, swarm intelligence inspires algorithms that enable multiple robots to perform tasks collectively without centralised control. This project leverages these principles, applying them to the realm of simulated combat, where each agent operates based on local information and interactions.

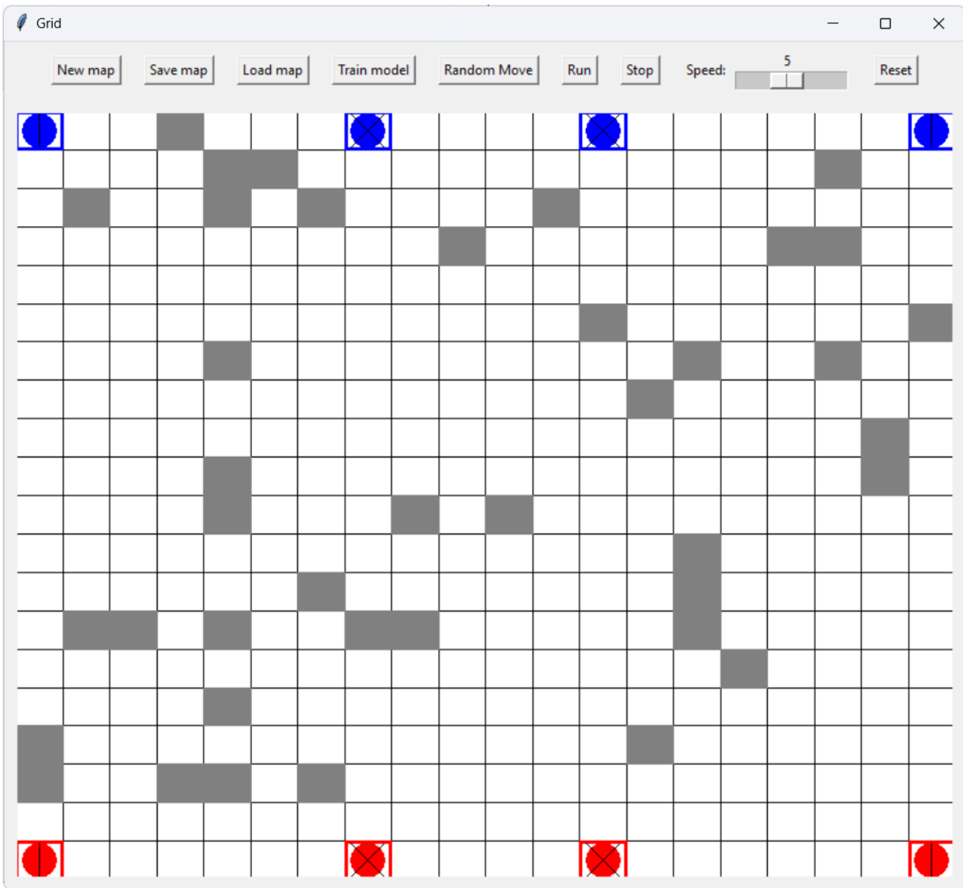

Figure 1: Application interface: *there are 2 opposing teams: blue and red, agents with symbol "X" are melee fighters and agents with symbol "|" are distance fighters.*

2 Methodology

- **Agent Design**: Agents are categorised into ranged and melee units, each with distinct combat strategies and capabilities. Neural networks dictate the behaviour of these agents, optimised through Q-learning. Each agent has a limited set of actions it can perform: move up, down, left, right, or attack.

- **State Representation**: Each agent perceives its environment through a state representation encompassing its immediate surroundings and health points. This ensures the agent's decisions are based on local information, promoting adaptability to different map configurations, similar to the localised behaviour seen in swarm intelligence systems.

- **Decision Making**: The decision-making process leverages DRL, where agents learn optimal strategies to outmanoeuvre and outlast the opposing team. The neural network's training is guided by Q-learning, adjusting actions to maximise cumulative rewards. Despite the simplicity of their actions, agents develop complex strategies through learning and interaction.

- **Framework Development**: A flexible framework was developed to support easy map generation and agent customisation. This enables users to experiment with various combat scenarios and agent configurations, fostering a deeper understanding of DRL in dynamic environments.

- **Application for Testing and Visualisation**: To facilitate testing and visualisation of the simulations, an intuitive application was created. This application provides a user-friendly interface for setting up experiments, running simulations, and visualising the results in real-time. It allows users to tweak parameters, generate maps, and observe agent behaviours dynamically, making it an invaluable tool for both development and demonstration purposes.

- **Fine Tuning** : In an effort to maximise the efficiency of our agents, we conducted extensive tests with various reward configurations and neural network architectures. This iterative process involved fine-tuning the reward values and network parameters to achieve a balanced training regimen, ensuring that agents developed robust and strategic behaviours. Additionally, we trained our agents in diverse environments, leveraging the locality of perceived information. This approach enabled our agents to adapt to a wide range of situations and scenarios. By experimenting with these variables, we successfully optimised the performance of our agents, resulting in more sophisticated and effective combat strategies.

3 Rewards system

To train the agents effectively, we implemented a reward structure designed to encourage valid actions, promote combat, and minimise negative behaviours. The reward structure is as follows:

1. **Movement Rewards**:

   - **Valid Move**: Agents receive a *really small reward* for making a valid move (up, down, left, right).
   - **Invalid Move**: Agents receive a *small penalty* if their movement attempt fails (e.g., moving into an obstacle or out of bounds).

   Objective: Encourage agents to perform valid moves and explore the environment efficiently.

2. **Attack Rewards**:

   - **Hit Enemy**: Agents receive a *high reward* for successfully hitting an enemy. This reward is multiplied by the number of enemies hit in a single attack.
   - **Hit Ally**: Agents receive a *big penalty* for hitting an ally. This penalty is multiplied by the number of allies hit in a single attack.
   - **Kill Enemy**: Agents receive an additional reward of *2 times the hit enemy reward* if their attack results in killing an enemy.

   Objective: Encourage agents to attack enemies effectively while avoiding friendly fire.

3. **Death**:

   - **Agent Death**: No negative reward is given if an agent dies during the simulation. This decision is made to encourage agents to engage in combat without the fear of penalisation for dying.

   Objective: Foster an aggressive and dynamic combat environment within the simulation.

4 Analysis

The project underscores the potential of DRL in simulating complex, dynamic environments where agents must rely on local information. By integrating principles of swarm intelligence, the framework and accompanying application significantly enhance the ease of experimenting with different scenarios and visualizing outcomes. Future enhancements could include incorporating more sophisticated state representations, such as temporal information, and exploring different learning algorithms to further refine agent behaviors.

Some interesting behaviors were observed:

- **Team up**: Agents from the same team take action together to eliminate an opponent.
- **Ambush**: Some agents do not move, but remain in ambush to eliminate any opponents who approach.

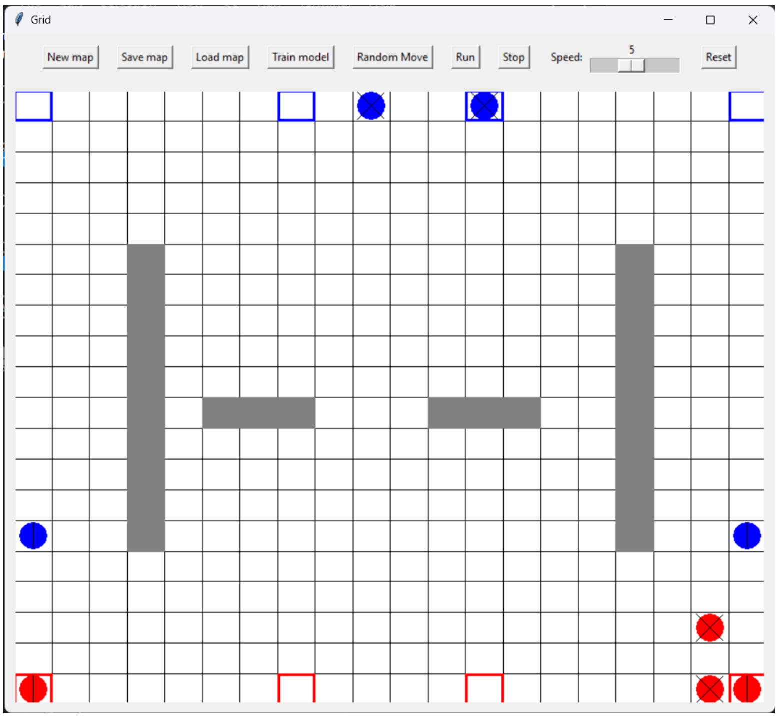

Figure 2: In simulation: 2 blue agents go on the attack while red agents wait for the opponent.

5 Conclusion

Swarm Warriors offers a robust, adaptable framework for simulating team-based combat scenarios using deep reinforcement learning. By focusing on local information and swarm intelligence principles, the project provides a scalable solution for studying coordinated agent behaviours in dynamic environments. The addition of an application for testing and visualisation further enriches the user experience, making the framework accessible and practical for a wide range of applications.