Prey-Predator Maze with Reinforcement Learning

Juan Diego Grajales Castillo
Proyecto Curricular Ingeniería de Sistemas
Universidad Distrital Francisco José de Caldas
Email: jdgrajalesc@udistrital.edu.co

Nicolás Castro Rivera
Proyecto Curricular Ingeniería de Sistemas
Universidad Distrital Francisco José de Caldas
Email: nicastro@udistrital.edu.co

***Abstract*—There is a complex system involving a prey and a predator inside a maze. Each has its own objective: to survive and to hunt, respectively. To model this system, we have designed a proposal with two cybernetic agents that use reinforcement learning to maximize their objectives in a competitive scenario. This project explores whether this model is effective through the use of AI and cybernetic foundations.**

***Index Terms*—Reinforcement Learning, Multi-Agent Systems, Prey-Predator, Cybernetics, Maze Simulation**

I. Introduction

To explore a system, it is necessary to apply systems thinking and evaluate the different components such as the prey, the predator, and the parts of the maze. This can be challenging if the system's sensitivity is not taken into account or if the model is poorly posed.

For this problem, it is important to treat the predator and the prey as adaptive cybernetic agents. Consequently, components such as maze parts, the presence of traps, the paths the agents can take, and the position of walls are fundamental for the modeling.

To address this problem, Game Theory can be used, as suggested in the paper: *Game Theory and Multi-Agent Reinforcement Learning: A Mathematical Overview* [1]. This article proposes using Game Theory to develop algorithms for strategic decision-making, including for autonomous agents in complex systems. It focuses on non-cooperative games and Nash equilibria, and extends the principles of single-agent reinforcement learning to multi-agent settings (MARL), particularly in fully adversarial games where each agent maximizes its benefit often at the expense of others.

This method can employ algorithms like Q-Learning, DQN, and Policy Gradient methods. It is a good fit for the prey-predator system, which aligns with the characteristics described in the paper.

Another perspective is presented in *Reinforcement Learning in Two Player Zero Sum Simultaneous Action Games* [2]. This paper focuses on zero-sum games, where two players use different strategies and compete for their objectives. In such games, one player's win is another's loss.

It also discusses concepts like Nash equilibria and algorithms like DQL, but contributes new ideas such as simultaneous actions and meta-learning. This is suitable for the prey-predator system, where each agent can earn rewards by hindering the other—making it a clear example of a zero-sum game: either the prey survives or the predator captures it.

Our proposed model uses reinforcement learning for multi-agent interaction.

II. Methods and Materials

In the first place, much time was devoted to analyze and design of the system. Identifying 5 main components:

- **Predator**: Agent, seeks to hunt prey.
- **Prey**: Agent, seeks to survive.
- **Traps**: Element that kills the prey when it passes over it
- **Walls**: Element that prevent the passage of agents
- **Path**: Where agents can transit.

Evaluating the different interrelations between these components is possible to define the synergy between them. The components like walls, traps, and the path are fixed in the environment, but prey and predatory are not. They will interact with elements and between them. The agents will receive the information using virtual proximity sensor, they detect the elements for take a decision for act with movement actuators. If any detects wall they wont be able to pass, if detect path will be able to move in that direction, if prey detect a trap or the predator, will change the direction of movement. In case of predator if detect the prey, will take direction to it for hunting.

HERE GO CYBERNETIC DIAGRAMS AND EXPLANATION OF LOOPS

REWARDS EXPLANATION

Each agent will have its own decision-making model based on reinforcement learning. The predator is trained to minimize the time it takes to capture the prey, while the prey is trained to maximize its survival time by avoiding both traps and the predator.

HERE GO DYNAMIC SYSTEM WITH MATHEMATICAL MODEL

Design decisions such as the reward structure, learning algorithm, and exploration strategy are crucial and must be carefully defined.

We are considering Q-Learning and Deep Q-Networks as the basis for training, depending on the complexity and dimensionality of the state space.

III. Results

IV. Conclusions

Acknowledgment

References

[1] M. Noguer and A. Mfougouon, "Game theory and multi-agent reinforcement learning: A mathematical overview," https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4926191, 2024.

[2] P. Phillips, “Reinforcement learning in two player zero sum simultaneous action games,” https://arxiv.org/abs/2110.04835, 2021.