Pacman Capture the Flag
Assignment 4 Report

GROUP 03

Shekhar Devm Upadhyay   Aiman Shenawa

20010531   001021

sdup@kth.se   ashenawa@kth.se

May 2023

Abstract

This report presents a study on the development of AI agents for controlling a team of two in a competitive Pac-Man Capture the Flag (CTF) scenario. The objective of the agents is to maximize their score by efficiently collecting food from the enemy turf within a limited number of moves. This report studies various approaches used in the field of artificial intelligence, including minimax game tree search, modifications like alpha-beta pruning and move ordering, Monte Carlo Tree Search (MCTS), Expectimax, and decision-tree-based approaches, presents the final solution employed by the authors, and does a comparative analysis of the approaches utilized by all the teams in the competition. By analyzing and comparing these strategies, we aim to identify effective techniques for AI-controlled agents in Pac-Man CTF.

1 Introduction

With the continious development of society and technology, the need for autonomous agents is increasing. It is becoming more and more important to develop agents that can operate in dynamic and complex environments autonomously. In this report we will explore the problem of developing an agent that can play the game of Pacman capture the flag. This is a project developed by UC Berkely [UC 15], it is a challenge that combines classic video game mechanics with artificial intelligence techniques. The goal of the project is to develop an agent that can play the game of Pacman capture the flag. The game is played on a grid, where the agent can move in four directions, up, down, left and right. The grid is split into two halves, in each half there is a team of Pacman agents, the goal of the game is to capture the food pellets from the other team's side and bring them back to your side while also preventing the other team from doing the same. Multiple aspects of artificial intelligence will be explored in this report, such as minimax game tree search, Monte Carlo Tree Search (MCTS), Expectimax, and decision-tree-based approaches. The final solution included the use of Minimax and alpha-beta pruning.
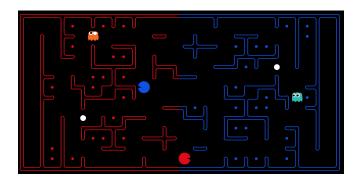

Figure 1: Illustration of environment

The figure shows the game environment, the blue and red dots represent the food pellets, the blue and red lines represent the walls. As seen in the figure above there are both Pacman and ghosts. The agent is a ghost when defending its own side. When crossing over to the enemy side to collect pellets it becomes a pacman. An agent can go into enemy territory and collect pellets, but the score will only increase when the agent returns back to its own side. In the figure there are also white pellets, those are power pellets, when an agent collects a power pellet it can eat the enemy agents for a period of 40 moves. During that period the enemy ghost agents are regarded as scared. When an agent is eaten it respawns at its own side, if the agent was scared before getting eaten it will no longer be scared after respawning. The game ends when all but two food pellets have been eaten, or when the time limit of 1200 moves has been reached. The team with the highest score wins the game. There is a limit on calculation time, 15 seconds for initialization and 1 second for each move. If the agent exceeds the time limit three times duting a game the team will forfeit the game.

The information each agent has access to is the following:

- The friend agent's positions

- The total score

- The position of all the food pellets on the map, this is updated every time an agent moves

- The position of all the power pellets on the map, this is updated every time an agent moves

- The position of an enemy agent if it is within 5 squares of the agent or a teammate

1.1 Task Description

This report will discuss the problem of developing an agent that can play the game of Pacman capture the flag. The aim is to develop two agents which are able to defend their side of the map and also collect food pellets from the enemy side. The agents will be able to move in four directions, up, down, left and right.

These problems and their solutions are addressed in further detail in the following section 2.

1.2 Contribution

In this report we will impement Minimax and alpha-beta pruning into the UC Berkeley Pacman Capture the Flag project. We will also explore the use of Monte Carlo Tree Search (MCTS), Expectimax, and decision-tree-based approaches. We will also compare the results of our solution with the results of other methods used for the same problem.

An essential aspect of Pacman capture the flag is the ability to cooperate as a team and communicate with the other agent. Potential real world applications of this project are many, as the ability of intelligent agents to make strategic decisions in a dynamic environment has many applications. Applications include multi-robot system coordination, autonomous vehicles, and unmanned aerial vehicles.

1.3 Outline

This report is organized as follows. First we introduce the problem and its importance in Section 1. Then we discuss and highlight the related work and required backround knowledge in order to understand the proposed solution in Section 2. This is followed by a detailed description of the proposed solution in Section 3. Next, we present the experimental setup and the results in Section 4. Finally, we summarize the results and conclude the report in Section 5.

2 Related Work

The success of AI agents in Pac-Man CTF heavily relies on the choice and implementation of appropriate strategies. This section provides an overview of the relevant research in this domain.

2.1 Minimax Game Tree Search

The problem of finding the optimal move in a game can be formulated as a search problem. The search space is a tree, where each node represents a state of the game, and each edge represents a possible move. The root node represents the current state of the game, and the leaf nodes represent the terminal states of the game. The goal is to find the optimal path from the root node to a leaf node.

The minimax algorithm [NM44] (shown in Algorithm 1) is a recursive algorithm that computes the optimal move for a player in a two-player game. The algorithm assumes that the opponent plays optimally, and it tries to minimize the maximum loss that can be incurred by the player. It is a classic approach in Game Theory, aiming to find the optimal move in a two-player, zero-sum game. By constructing a game tree and evaluating the utility of different game states, the algorithm enables agents to make informed decisions.

Sometimes, however, it is not feasible to search the entire game tree, due to the large number of possible moves. To address this issue, the algorithm can be modified to limit the depth of the search tree. The algorithm can also be modified to incorporate heuristics that evaluate the utility of non-terminal game states. Variations of minimax, such as alpha-beta pruning, move ordering based on heuristics, iterative deepening search (IDS), and Repeated States Checking (RSC), have been proposed to improve its efficiency and effectiveness [CHH02].

Algorithm 1 Minimax Algorithm

1: **procedure** MINIMAX($node, depth, maximizingPlayer$)
2:      **if** $depth = 0$ or $node$ is a terminal node **then**
3:          **return** the heuristic value of $node$
4:      **end if**
5:      **if** $maximizingPlayer$ **then**
6:          $value \leftarrow -\infty$
7:          **for** each child of $node$ **do**
8:              $value \leftarrow \max(value, \text{Minimax}(child, depth - 1, \text{False}))$
9:          **end for**
10:         **return** $value$
11:      **else**
12:         $value \leftarrow \infty$
13:         **for** each child of $node$ **do**
14:             $value \leftarrow \min(value, \text{Minimax}(child, depth - 1, \text{True}))$
15:         **end for**
16:         **return** $value$
17:      **end if**
18: **end procedure**

2.2   Expectimax

Expectimax is a variant of the minimax algorithm that considers uncertain outcomes in games. It is commonly used in domains with probabilistic elements. In Pac-Man CTF, Expectimax can be employed to account for the ghost movement and uncertain states, enabling agents to make rational decisions under uncertainty [Riv87].

2.3   Monte Carlo Tree Search

MCTS is a sampling-based search algorithm that has demonstrated remarkable success in various game-playing domains. By combining tree exploration with random rollouts, MCTS performs effective exploration and exploitation of the search space. This approach has been applied to

Pac-Man CTF, where agents simulate multiple playouts to estimate the value of different actions and make informed decisions [GS11].

2.4 Decision Tree Based Approaches

Decision trees offer a structured representation of decision-making processes. By learning from historical data and constructing decision rules, decision-tree-based approaches provide an interpretable framework for agent behavior. These techniques have been explored in Pac-Man CTF, where agents use decision trees to guide their actions based on various game state features [HTF09].

3 Method

In this section, we will discuss the solution proposed by our team. The solution is based on the Minimax algorithm, with some modifications to account for the limitations of the UC Berkeley Pacman Capture the Flag project. We will also discuss the motivations behind our heuristic functions.

3.1 Modified Minimax

Typical Minimax approaches require that the entire game state be observable to us. Firstly, the UC Berkeley challenge does not provide us with the entire game state - we can only see the positions of our team members, and nearby enemies. Secondly, even if we knew the positions of all the enemies, if we consider the movements of both agents of our team as a single action, then the branching factor of the game tree is too large to be explored in a reasonable amount of time. So, we decided to modify the typical Minimax approach as follows:

- At each time step, we consider the possible actions of the agent we are controlling (as the maximizing player). If there is an enemy agent visible to us, we consider the possible actions of the enemy agent as well, by passing control to the enemy as a minimizing player. If not, we continue to consider the possible actions of our agent (by exploring further based on moves that we can make in the next time step as the maximizing player).

- We use a heuristic to evaluate the utility of a game state, instead of evaluating the utility of a terminal state. We compute heuristics until a certain depth of the game tree, and use iteratively deepening search to explore the game tree further if time permits.

- We use alpha-beta pruning to reduce the number of nodes that we need to explore in the minimax exploration of the game tree. To further speed up the search, we use move ordering based on heuristics to explore the most promising nodes first, and implement repeated states checking to avoid computing values of the same game state multiple times when working with deeper game trees.

3.2 Heuristics

We started off by having separate classes for offensive and defensive agents, each with a different heuristic. Later, we merged the two classes into a single class, and used a attribute called `mode`

to determine if the behavior should be defensive or offensive, based on current agent position, position of our teammate, positions of any visible enemies, etc.

3.3   Offensive Heuristics

When in "offensive" mode, our agent tries to cross over to the other side of the map as quickly as possible, then actively looks for food pellets to capture. We prioritize closer food (without disregarding the rest), de-value food inside dead-end paths (paths that can be choked off at a single point by the enemy) pick up power pills when possible and when enemy agents are closeby. As the food in our pocket increases, we have an increasing tendency to return to our side of the map to deposit the food. We also try to avoid enemy agents when possible, and try to eat them when they are scared. Being chased by an unscared enemy also "turns off" the tendency to chase food, and we try to run away from the enemy instead. We also try to avoid getting trapped in dead-end paths, and try to avoid getting cornered by enemy agents. After depositing the food, we start this process again.

We noticed that the offensive agents sometimes got stuck in local extrema of the value function. For example, sometimes the agent sat next to a food particle and didn't eat it, until the enemy agent came closer and pushed it into scramble mode. To avoid this, we added an increasing tendency to return to our own side for each time step in which the chosen action was to stay in the same place. This helped the agent get out of local extrema in a couple of moves, in most cases.

3.4   Defensive Heuristics

For the "defense" mode of the agent, we incentivize minimizing the sum of distances to all the food we are defending. Since food near the center line is easier to take by the enemy, we give more weight to being close to those food particles. We keep track of the food that is being defended by our agent, and if some of it disappears, we know that an enemy agent has eaten it. We try to chase the enemy agent based on direct sight, knowledge from our teammate, and disappearing food. If we are close enough to the enemy agent, we move to trap and eat it we are not scared. If we are scared, we try to run away from the enemy agent. In no circumstance does the agent go over to enemy territory when in defense mode.

The exact details of our implementation can be found in our Github Repository and submitted final code.

4   Experiments and Results

First, we will describe the experimental setup for the two problems, and then we will present the results for each of the problems separately.

4.1 Experimental Setup

4.1.1 Experimental setup for measurement of performance

In order to evaluate the performance of the suggested solution, we will compare our approach with five other groups. This is done by putting the algorithms against each other in the simulation. Two different evaluations will be done, one called pre-finals, in which the groups give their initial solutions, and one called finals, in which after having time to re-evaluate their solutions, the groups give their final solutions. The results of both evaluations are presented in the tables below. Results are measured in terms of wins, ties, losses. The score is calculated as the following: 3 points for a win, 1 point for a tie, and 0 points for a loss. The total score is the sum of the points for each group.

4.2 Results

In this section, we summarize the results of our experiments. First we present the results of the two evaluations in the competition, then we present the approach of the winning group.

Group	Wins	Ties	Losses	Total Score
4	8	0	0	24
2	5	1	0	16
6	5	1	3	16
3	3	5	2	14

Table 1: Results of the pre-finals evaluation.

Group	Wins	Ties	Losses	Total Score
4	13	1	1	40
2	9	2	4	29
6	7	2	6	23
3	4	4	6	16

Table 2: Results of the finals evaluation.

4.3 Improvements

In this section we will discuss the improvements that can be made to the suggested solution.

During the development of the solution, we noticed that the agents sometimes got stuck in local extrema of the value function. For example, sometimes the agent sat next to a food particle and didn't eat it, until the enemy agent came closer and pushed it into scramble mode. To avoid this, we added an increasing tendency to return to our own side for each time step in which the chosen action was to stay in the same place. This helped the agent get out of local extrema in a couple of moves, in most cases. However this is not a perfect solution, as we risk getting stuck in a position within a chokehold in the enemy territory. This roots to the heuristic design we have chosen and in order to improve upon this a in depth analysis of the different heuristics is required.

In the finals we observed a peculiar behavior of the offensive agent, as the enemy ghost

5 Summary and Conclusions

References

[CHH02] Murray Campbell, A.Joseph Hoane, and Feng-hsiung Hsu. "Deep Blue". In: *Artificial Intelligence* 134.1 (2002), pp. 57–83. ISSN: 0004-3702. DOI: `https://doi.org/10.1016/S0004-3702(01)00129-1`. URL: `https://www.sciencedirect.com/science/article/pii/S0004370201001291`.

[GS11] Sylvain Gelly and David Silver. "Monte-Carlo tree search and rapid action value estimation in computer Go". In: *Artificial Intelligence* 175.11 (2011), pp. 1856–1875. ISSN: 0004-3702. DOI: `https://doi.org/10.1016/j.artint.2011.03.007`. URL: `https://www.sciencedirect.com/science/article/pii/S000437021100052X`.

[HTF09] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning.* New York, NY: Springer, 2009. ISBN: 978-0-387-84857-0.

[NM44] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press, 1944.

[Riv87] Ronald L. Rivest. "Game tree searching by min/max approximation". In: *Artificial Intelligence* 34.1 (1987), pp. 77–96. ISSN: 0004-3702. DOI: `https://doi.org/10.1016/0004-3702(87)90004-X`. URL: `https://www.sciencedirect.com/science/article/pii/000437028790004X`.

[UC 15] UC Berkeley. *Contest: Pacman Capture the Flag.* `http://ai.berkeley.edu/contest.html`. 2015.

A Appendix