James Soto

4-9-2023

CS-370 Current and Emerging Trends

Project Two – Design Defense

Analyze the differences between human and machine approaches to solving problems.

Describe the steps a human being would take to solve this maze.

First a human will observe the maze and identify as much information as possible, like the starting point and the layout. Depending on the individual they could utilize multiple methods to solve the maze. One method is trial and error where the human will walk down a path, making decisions based on his/her intuition and observation. If they hit a dead end the human will retrace their steps and try a different path until they reach the end/goal.

Describe the steps your intelligent agent is taking to solve this pathfinding problem.

In Q-learning, the intelligent agent(pirate) starts with zero knowledge of the environment(maze) and its objective inside the maze or what actions will lead to rewards or punishments. As the intelligent agent(pirate) plays the game, it quickly learns the reward system and realizes that it is getting punished for wandering or taking paths that do not lead to a reward. The intelligent agent(pirate) saves its experience from the actions it has taken in a particular state and stores that experience as a numerical value in its Q-table. The intelligent agent(pirate) keeps track of the visited nodes and their respective distances from the starting node. The values in the Q-table indicate how likely the intelligent agent(pirate) is to take that action again in the future. The intelligent agent(pirate) identifies when it has reached the treasure (end node or end of the maze) and stops it’s search. From here the intelligent agent(pirate) reconstructs the optimal path from the start to the end nodes, using the values from the Q-table experience.

What are the similarities and differences between these two approaches?

Similarities:

Both approaches involve a process of exploration and navigation through the maze.

Both will involve some trial and error, where they test different paths until the optimal solution is found.

Both can adapt and learn, improving their problem-solving capabilities, humans will gain from experience and develop new strategies while machines use RL for training to optimize performance.

Differences:

Human problem-solving relies on observation, pattern recognition, intuition, and experience, while machines use data structures, algorithms, and computational power.

Machines can systematically explore and evaluate all possible paths in a maze while processing large amounts of data in a couple minutes, while humans may be more limited by memory and cognitive resources while taking longer to solve the problem.

Humans could make informed decisions, while machines are limited to their decisions by the training data and rules that were provided.

Assess the purpose of the intelligent agent in pathfinding.

What is the difference between exploitation and exploration? What is the ideal proportion of exploitation and exploration for this pathfinding problem? Explain your reasoning.

From the reading throughout this course, my understanding of the purpose of the intelligent agent in pathfinding is to efficiently find the shortest or most optimal path from a starting point to the end or goal point in a given environment, which for our project was a maze. The difference between exploitation and exploration is that exploitation is the process of the agent utilizing existing knowledge and existing experience to make its decisions, while focusing on the actions that have proven to be rewarding. Exploration refers to the process of gathering new information or data by trying new actions or taking new paths, without being impacted by rewards, looking for the potentially more effective routes/paths than the currently known paths. For this pathfinding example the ideal proportion would be a balance between exploration and exploitation to ensure efficient learning and decision-making. During exploration the agent gathers the information about potential paths, and while the agent gains more knowledge about the paths, and environments, exploitation becomes more important to optimize the solution based on the acquired information the agent gathered during exploration.

How can reinforcement learning help to determine the path to the goal (the treasure) by the agent (the pirate)?

The agent can use RL to determine the most efficient path to the goal(treasure). The agent represents the environment as a set of states transitioning based on available actions like where to move. The agent then chooses actions based on the policy, which initially the policy could be random or predefined, to map from states to actions. The agent then receives a reward, or a penalty based on its actions. As the agent learns based on the rewards it has received it will update the policy. Ultimately with repeated interactions with the environment and receiving a reward or penalty and then iteratively updating its policy, the agent learns the most efficient and effective way to the treasure.

Evaluate the use of algorithms to solve complex problems.

How did you implement deep Q-learning using neural networks for this game?

The objective is to teach the pirate agent how to find the treasure by taking the best actions in the game environment. The agent must learn to navigate the grid while avoiding obstacles and reaching the treasure. The implementation I used for the training loop for deep Q-learning using NN was that for each episode the environment would reset, and the state would be initialized. The agent randomly selects a free cell. The agent explores with probability epsilon and exploits the current Q-values. The agent executes the chosen action and observes the new state and reward. The agent stores the experience (state, action, reward, next state, done). The NN gets trained by using the experience data and the target Q-values. In my instance I performed multiple tests changing epochs and batch size to find the most optimal solution to the problem. For this training I found that epochs 10 and batch size 32 reached 100% win rate at epoch 115, which was the lowest total from all my testing.