Pirate Maze Design Defense

Jon Wickerd

10/19/2024

**1. Analyze the Differences Between Human and Machine Approaches to Solving Problems**

**Human Approach**

Humans often solve problems using a problem-solving approach, this means they rely on intuitive strategies and past experiences. For instance, when faced with a maze, a person might visualize the layout and plan a path based on what has worked in the past. Typically, a human would look for open areas, avoid dead ends, and backtrack when they encounter obstacles. This process is informed by memory; previous experiences with similar mazes can help individuals make better decisions in new situations.

**Machine Approach**

In contrast, a machine similar to the pirate in this project, learns to navigate the maze through trial and error using a mathematical model. It systematically explores various paths and keeps track of the rewards or penalties associated with each action taken. This learning process is facilitated by reinforcement learning, this allows the agent to predict the value of future actions based on its past experiences (Mnih et al., 2015).

**Similarities and Differences**

Both humans and machines share similarities in their problem-solving approaches. For instance, both utilize prior experiences to inform future decisions and explore different paths before identifying the optimal solution. However, there are notable differences. Humans are generally better at generalizing from previous knowledge and can apply reasoning to solve problems, while the machine relies solely on numerical rewards and systematic exploration of possibilities without any intuitive understanding (Russell & Norvig, 2020).

**2. Assess the Purpose of the Intelligent Agent in Pathfinding**

The primary purpose of the intelligent agent in pathfinding is to identify the optimal route from its starting position to the goal, in this is project, it is the treasure. Utilizing deep Q-learning, the agent continuously updates its understanding of the environment based on the rewards it receives from various actions taken during its exploration (Mnih et al., 2015). The overarching goal is to maximize cumulative rewards, which directs the agent towards the most efficient pathway to the treasure.

**Exploitation vs. Exploration:**

Exploitation refers to the strategy where the agent chooses actions based on its current knowledge, aiming to maximize rewards. In this phase, the agent selects the path it believes to be the best based on its past experiences.

Exploration, on the other hand, allows the agent to try out different actions, potentially discovering new paths that were previously unknown. This phase is essential for updating the agent's knowledge of the environment.

**Ideal Proportion**

An effective balance between exploitation and exploration is often achieved by setting the exploration factor (epsilon) around 0.1. This means the agent will engage in exploration 10% of the time and rely on exploitation 90% of the time. This proportion can be adjusted depending on the learning speed of the agent; for instance, as the agent becomes more proficient, reducing exploration may yield better results (Sutton & Barto, 2018).

**Reinforcement Learning’s Role:**

Reinforcement learning plays a crucial role in guiding the agent towards the treasure. The agent receives rewards for reaching specific states, with the highest reward assigned for successfully arriving at the treasure. By leveraging these rewards, the agent learns which actions lead to positive outcomes, allowing it to refine its strategy over time (Mnih et al., 2015).

**3. Evaluate the Use of Algorithms to Solve Complex Problems**

Deep Q-Learning is a powerful algorithm utilized to address complex pathfinding challenges, such as navigating through a maze in our treasure hunt game. In this implementation, we designed a neural network that estimates Q-values for each possible action based on the current state of the environment. The architecture consists of an input layer that matches the maze size (an 8x8 matrix), multiple hidden layers using PReLU activation functions, and an output layer that generates Q-values for actions defined as moving left, up, right, or down.

To enhance the stability and efficiency of the training process, our algorithm employs a technique known as experience replay. This method involves storing past experiences—comprising the state, action, reward, and subsequent state—in a memory buffer, managed by the GameExperience class provided in the starter code. During training, the neural network is updated using mini-batches of these experiences, which helps mitigate the correlation between consecutive experiences and facilitates more effective learning.

Furthermore, the agent adopts an epsilon-greedy strategy for action selection, where it explores new actions 10% of the time (as defined by the exploration factor, epsilon). This allows the agent to discover potentially better paths while primarily leveraging known actions to maximize its rewards.

Overall, Deep Q-Learning integrates neural networks with reinforcement learning principles, creating a robust framework capable of efficiently solving intricate pathfinding problems, as evidenced in our pirate agent’s quest to find the treasure before the human player (Mnih et al., 2015).

References:

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.

Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach. Pearson.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press.