Victor Udeh

CS370 Module 7

Aug 12th, 2024

7-3 Project Two

Design Defense

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

**Human Approach:** When a human attempts to solve a maze, they rely heavily on visual and spatial reasoning. A human might start at the entrance and attempt to find the exit by choosing a direction and following it, remembering paths that lead to dead ends and backtracking when necessary. Humans often use trial and error, heuristics (such as always turning right), or even intuition based on previous experiences. This approach can be time-consuming, especially in larger or more complex mazes, and humans may not always find the optimal path.

**Machine Approach:** The machine, in this case, the pirate intelligent agent, solves the maze using a deep Q-learning algorithm. Unlike humans, the agent does not rely on visual cues or spatial reasoning but rather learns from interactions with the environment. The agent explores the maze by taking actions and receiving feedback in the form of rewards or penalties. Over time, the agent learns to associate certain states (positions in the maze) with higher expected rewards, which guide its decision-making process to find the most efficient path to the treasure.

**Similarities and Differences:** Both human and machine approaches involve exploration and learning from mistakes. However, the human approach is more intuitive and may involve shortcuts based on visual or prior knowledge. In contrast, the machine relies on mathematical models and systematic exploration to learn the optimal path. The machine can handle much more complex environments and find optimal solutions more reliably, while humans might be faster in simple scenarios but less effective in more complex ones.

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

The purpose of the intelligent agent in this pathfinding problem is to autonomously navigate the maze to find the treasure before the human player does. The agent must learn to identify the shortest and safest path while avoiding obstacles and dead ends. The effectiveness of the agent is determined by how well it balances exploration (trying new paths) and exploitation (using known paths that have proven to be successful).

**Exploitation vs. Exploration:**

* **Exploitation** involves the agent using the knowledge it has already gained to make decisions that maximize immediate rewards. In this context, it means following paths that are known to lead closer to the treasure.
* **Exploration** involves trying out new or less certain paths to discover whether they might lead to better rewards in the long run.

**Ideal Proportion:** The ideal proportion of exploitation to exploration depends on the specific problem and the complexity of the maze. Initially, a higher rate of exploration (e.g., an epsilon of 0.1) is important to allow the agent to learn about the environment. As the agent becomes more knowledgeable, it should gradually increase exploitation to refine its pathfinding strategy and avoid unnecessary risks. A common approach is to start with a high exploration rate and then reduce it as the agent's performance improves, allowing for more exploitation of the learned strategies.

**Reinforcement Learning in Pathfinding:** Reinforcement learning helps the agent learn the best path by associating actions with rewards. When the agent takes an action that brings it closer to the treasure, it receives a positive reward, reinforcing that behavior. Conversely, actions that lead to dead ends or further away from the treasure result in penalties. Over time, the agent learns which actions maximize rewards, effectively determining the optimal path to the treasure.

**3. Evaluation of Deep Q-Learning for Complex Problem Solving**

**Implementation of Deep Q-Learning:** Deep Q-learning is implemented using a neural network that approximates the Q-value function. This function predicts the expected future rewards for each possible action in a given state. The network is trained using experiences collected during the agent's interactions with the environment, stored in a replay memory. By sampling from this memory and updating the network's weights, the agent learns to predict the Q-values more accurately, allowing it to make better decisions over time.

* **Neural Network Architecture:** The network used in this game consists of three layers. The input layer corresponds to the current state of the maze, while the output layer provides Q-values for each possible action (left, right, up, down). The hidden layers allow the network to learn complex relationships between states and actions.
* **Training Process:** The agent trains over multiple epochs, with each epoch representing an attempt to navigate the maze. The loss function, typically mean squared error (MSE), measures the difference between predicted Q-values and the target values, which are updated based on the rewards received during training.
* **Challenges:** Some challenges in implementing deep Q-learning include managing the exploration-exploitation trade-off and ensuring that the agent does not get stuck in local optima. The use of experience replay helps to mitigate these issues by breaking correlations between consecutive experiences.

Deep Q-learning is well-suited for this type of problem because it allows the agent to learn from experience in a way that generalizes well to unseen scenarios. The neural network’s ability to approximate complex functions makes it an effective tool for pathfinding in environments like the maze in this treasure hunt game.

This design defense outlines the approach taken in developing the pirate intelligent agent, highlighting the differences between human and machine problem-solving, the role of exploration and exploitation in reinforcement learning, and the implementation of deep Q-learning for this pathfinding task. By utilizing these techniques, the intelligent agent can autonomously and efficiently navigate the maze, providing a competitive challenge for the human player.

**References**

* Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.
* Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529-533.
* Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.