Dennis Ward

10/07/2024

CS370

Module 7-2

Defense Design

**Human and Machine Approaches to Solving Problems**

Humans and machines approach problem-solving in notably different ways, especially in pathfinding tasks. A human, when presented with the maze in this project, would generally start by analyzing the environment visually. They may consider prior knowledge of maze structures, use trial and error, and attempt to logically deduce the best path to reach the treasure while avoiding dead-ends. A person would likely make use of heuristics, shortcuts, and spatial awareness to minimize unnecessary exploration. This approach involves flexible thinking and intuition, as humans dynamically adjust their path based on immediate feedback.

In contrast, the machine, particularly the intelligent agent designed for this project, employs a deep Q-learning algorithm to solve the pathfinding problem. The agent navigates through the environment by taking actions (left, right, up, down), receiving rewards based on its actions, and learning from these rewards to optimize future behavior. The machine uses reinforcement learning to map its actions to rewards, building a model that maximizes cumulative future rewards by finding the optimal policy (Silver, 2016).

While humans rely on intuition and shortcuts, the machine follows a methodical trial-and-error approach through a combination of exploration and exploitation. In this implementation, the machine’s exploration rate was kept constant at 0.1, meaning 10% of the time, it performed random exploratory actions, and the remaining 90% was focused on exploitation of learned knowledge. This fixed exploration rate allowed the agent to balance discovering new paths and optimizing learned actions. Unlike humans, the machine lacks adaptability beyond its predefined algorithm, but it benefits from a consistent approach to learning (Hassabis, 2017).

**Steps in Solving the Maze**

A human would approach the maze by first scanning for a feasible path and noting obstacles. They would likely attempt a direct route toward the treasure, adjusting based on dead-ends or blocked paths. Humans would naturally avoid previously visited dead-ends and make decisions influenced by prior experience and cognitive shortcuts.

In comparison, the intelligent agent follows a series of steps dictated by the deep Q-learning algorithm. Initially, the agent randomly selects actions, moving through the maze and recording the rewards associated with each action. Since exploration remained fixed at 10%, the agent consistently explored 10% of the time, while exploiting its learned knowledge in 90% of the steps. This exploration allowed it to gather information on less traveled or unexplored paths, while exploitation enabled the agent to take actions it predicted would result in the highest reward based on its model. The agent trains a neural network to predict future rewards for each action and updates its policy as more experience is gathered (Mnih, 2015).

Both humans and machines adjust their behavior based on feedback from the environment. However, the key difference is that while a human may rely on intuition and contextual awareness, the machine strictly adheres to its algorithm, learning through the reward structure and optimization process.

**Purpose of the Intelligent Agent in Pathfinding**

The intelligent agent's primary purpose is to optimize its path to the treasure by learning from its environment. Reinforcement learning, particularly deep Q-learning, provides the agent with the ability to learn through trial and error and adjust its strategy over time. In this project, exploration was set to a constant value of 0.1, meaning that in 10% of the steps, the agent explored random paths, and in the remaining 90%, it exploited the knowledge it had learned to optimize its actions.

The decision to maintain a fixed exploration rate ensured that the agent could continue to discover new paths throughout the learning process without over-exploring or getting stuck in suboptimal areas. The exploration rate, in this case, allowed for sufficient environmental discovery while focusing primarily on learned behaviors to navigate the maze more efficiently (Sutton & Barto, 2018).

Reinforcement learning helps the agent determine the path to the goal by enabling it to understand the long-term value of its actions. Q-learning allows the agent to estimate the expected cumulative rewards from each action and state. Through repeated trials and feedback, the agent improves its ability to navigate toward the treasure, prioritizing actions that lead to higher rewards over time.

**Deep Q-Learning and Neural Networks in the Game**

To solve the treasure hunt problem, deep Q-learning was implemented using a neural network. The network takes the state of the maze as input and outputs the expected reward for each possible action. During training, the agent collects experiences comprised of the current state, the action taken, the resulting reward, and the new state. These experiences are stored in memory and used to train the neural network. The network learns to predict Q-values for each action, which represent the long-term expected rewards.

The model architecture includes multiple hidden layers, which enable it to learn complex relationships between the states of the maze and the expected rewards for each action. With every episode, the agent updates its policy by using these experiences to refine the neural network’s predictions. This allows the agent to continuously improve its ability to find the optimal path to the treasure over time (Mnih, 2015).

The fixed exploration rate of 0.1 was chosen to provide the agent with enough flexibility to discover new paths without causing excessive randomness, making the learning process more efficient. Deep Q-learning is particularly well-suited for this type of problem because it enables the agent to learn optimal strategies in environments with complex state spaces, such as the maze. The neural network empowers the agent to generalize across different states, allowing it to develop strategies that apply to a wide variety of situations within the maze.

References

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron, 95*(2), 245–258. <https://doi.org/10.1016/j.neuron.2017.06.011>

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. <https://doi.org/10.1038/nature14236>

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., … & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*(7587), 484-489. <https://doi.org/10.1038/nature16961>

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.