Design Defense

SNHU

Elora Newcomb

12/03/2023

**Introduction**

This task explores the differences between human and machine problem-solving approaches. Specifically, it looks at a maze-solving task and examines how an intelligent agent (which is referred to as the pirate for the project) would resolve the problem compared to a human being. In this milestone, we will discuss how the project was developed, and the similarities and differences to humans regarding how the pirate accomplishes the task in pathfinding.

The way in which humans and robots solve problems differs significantly. To find the prize in this maze, for example, a person would have to create a mental image of the maze first before they can figure out the most efficient route. While navigating said maze, they would potentially make use of familiar locations and hints, adopting a process of trial and error to determine the best solution (Mnih, et al., 2015). The pirate, on the other hand, would adopt a different approach; reinforcement learning can be used to determine the most efficient way to go after the treasure, allowing the pirate to implement experience replay to store and learn from its prior experiences. Following a random start in the labyrinth, it could use the Q-learning method to determine the most appropriate course of action.

Although both human and non-human strategies seek to find the treasure at the end of the maze, the human method’s approach is much more fluid and responsive to the maze's ever-shifting conditions, relying mostly on intuition and experience, as opposed to the machine's reliance on learning and replaying past iterations of the treasure hunt. To maximize its reward, the pirate first needs to learn the fastest and safest way to its goal (the treasure) by employing exploratory and exploitative strategies. The idea of exploration entails carrying out random activities to investigate the surroundings and gain knowledge, while exploitation involves acting in the best manner feasible under the circumstances.

Both the complexity of the maze and the number of possible courses of action contribute to determining the best strategy for overcoming the pathfinding problem. In the early stages of the pirate’s development, investigation focused on gathering as much data as possible (Sutton & Barto, 2018), and as the agent gained knowledge about its surroundings, it became increasingly more engaged in exploitative behavior to maximize its reward. To illustrate this scenario, in the study "Playing Atari with Deep Reinforcement Learning" (Kilcher, 2020) stated that a 10% to 20% exploration rate is sufficient for most game, which proved to be true in this case with a 0.1% rate.

With the use of a reward signal, reinforcement learning assisted the agent in determining how to proceed. If they succeeded in completing a task, they would be rewarded; if they failed, they would be punished (Russell & Norvig, 2016). When making decisions and learning the best course of action, the agent was programmed to take such factors into account.

The agent trained the neural network by backpropagation after updating the Q-values with the Bellman equation. The Epsilon-greedy strategy, which strikes a balance between exploration and exploitation, was used to guide the agent's action selection. The neural network was then used to make decisions in real-time while the agent continuously learned and refined the network, allowing it to make better decisions over time.

In conclusion, the use of neural networks for deep Q-learning provides a powerful method for tackling difficult problems. Reinforcement learning allows the agent to alter its behaviors based on the reward signal, thus enabling it to discover the ideal approach to the objective through a combination of exploration and exploitation. Using a neural network, an intelligent agent, like the pirate in this exercise, can estimate the Q-values of its actions in each state, picking the action with the highest Q-value to maximize its payoff in the end.

**References**

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

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

Kilcher, Y. (2020, July 26). *Playing Atari with deep reinforcement learning (paper explained)*. YouTube. https://youtu.be/rFwQDDbYTm4?si=KjdqiXmjbW-496\_b&t=2032. 33:52. From: Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602. https://www.cs.toronto.edu/~vmnih/docs/dqn.pdf.

Russell, S. J., & Norvig, P. (2016). Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited.