**7-3 Project Two**

Karina Aronov

Bachelor of Science in Computer Science, SNHU

CS 370: Current/Emerging Trends in CS

Ramsey Kraya

February 23, 2025

**7-3 Project Two**

AI and humans have entirely different approaches to problem-solving. When trying to get through a maze, a human would use pattern recognition, instinct, and prior knowledge to determine the best path (Zhao & Marquez, 2013). They may begin by mentally outlining the layout, avoiding obstacles, and making decisions based on reasoning. They retrace their steps and try a different route if they become lost. Also, humans can improve their journey by using shortcuts like memory or visual clues.

On the other hand, my AI agent, the pirate, learns through reinforcement learning rather than relying on instinct. It starts by making completely random moves and slowly improves by learning from the rewards and penalties it receives. Every time it finds a successful path, it remembers it, reinforcing good choices while avoiding bad ones. Unlike humans, it doesn’t get frustrated or distracted, and it can attempt thousands of attempts in a short period, which makes it much more efficient at finding the best path (Engström, 2021).

The AI pirate solves the maze using Deep Q-Learning, a reinforcement learning technique that enables it to learn from experience. First, it explores the environment by making random moves, gathering data on possible paths and obstacles. This exploration phase allows it to understand the maze without prior knowledge. As it moves, it tracks its position and available actions, updating its internal system of the maze (GeeksforGeeks, 2023b).

As the training progresses, the AI shifts from exploration to exploitation. It starts using what it has learned to make better decisions rather than just randomly testing new paths. Using a deep neural network, the AI estimates the best possible move based on past experiences, helping it efficiently navigate toward the treasure while avoiding unnecessary moves (Mnih et al., 2015). This approach allows it to generalize and find the best solution more efficiently than traditional Q-learning.

While both humans and AI try to find the most efficient route, they go about it in very different ways. Humans tend to rely on reasoning and memory, sometimes making mistakes due to cognitive biases. AI, on the other hand, learns from trial and error and has no prior preconceptions (Botvinick et al., 2019). However, improper training might also cause the AI to struggle. If it doesn't conduct enough exploring, it may choose a poor course and never find a better one.

In reinforcement learning, exploration is when the AI tries new paths to discover better solutions, while exploitation is when it uses what it has already learned to make the best move. The right balance between these two is important. If AI only exploits what it knows, it might get stuck in an unsuccessful path. If it only explores, it wastes time testing unnecessary routes (GeeksforGeeks, 2023a).

For this treasure hunt game, a gradual shift from exploration to exploitation works best. At the start, the AI should explore more with a higher epsilon value, trying different paths to learn what works. Over time, it should reduce exploration and focus on using the most efficient path lowering epsilon. This ensures that the AI learns quickly but also settles on the best solution without wasting time.

Reinforcement learning is perfect for this problem because it allows the pirate to learn through experience. Instead of being directly coded with a path, it figures it out by making decisions and receiving feedback. This mimics how humans learn by trial and error but at a much bigger scale. Each move affects future rewards, so the AI must learn to plan ahead rather than just reacting to the current situation. Over time, it finds the shortest and safest way to the treasure without unnecessary moves (Silver et al., 2017).

The same idea is applied in real-world AI applications, such as robotics and self-driving cars, where AI must constantly react to changing circumstances. In the same way that a pirate figures out the best path through a maze, autonomous cars use their prior driving experiences to determine the safest and quickest route to navigate the roads (Botvinick et al., 2019). By continuously improving their decision-making, these AI-driven systems become more reliable and adaptable, making them useful in complex, real-world scenarios.

Deep Q-Learning is the main algorithm behind this game, allowing the AI to learn and optimize its strategy to find the path. It uses a deep neural network with multiple layers to process information about the maze, enabling it to make better decisions. Instead of storing every possible state-action pair, the AI relies on Q-Table approximation, using the neural network to estimate the best action based on its current position. Additionally, the AI utilizes experience replay, where it stores past experiences and reuses them to improve its decision-making process, leading to faster learning and more stable results. Over multiple training runs, the neural network gradually strengthens its ability to predict the most efficient moves. This method allows the AI to expand and identify the best solutions more effectively than traditional Q-learning, which has trouble with large state spaces (Mnih et al., 2015).

In the end, the AI pirate successfully learns to navigate the maze using reinforcement learning while humans rely on instinct and past experiences, the AI thoroughly tests paths, learns from rewards, and improves its strategy. By balancing exploration and exploitation, it efficiently discovers the best route to the treasure. This type of learning is widely used in AI applications, proving how reinforcement learning can solve complex problems beyond just games.

**Resources:**

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement learning, fast and slow. *Trends in Cognitive Sciences*, *23*(5), 408–422. https://doi.org/10.1016/j.tics.2019.02.006

Engström, M. (2021, December 8). Building an AI to navigate a maze - Magnus Engström - Medium. *Medium*. https://magnus-engstrom.medium.com/building-an-ai-to-navigate-a-maze-899bf03f224d

GeeksforGeeks. (2023a, January 10). *EpsilonGreedy Algorithm in Reinforcement learning*. GeeksforGeeks. https://www.geeksforgeeks.org/epsilon-greedy-algorithm-in-reinforcement-learning/

GeeksforGeeks. (2023b, January 23). *Deep QLearning*. GeeksforGeeks. https://www.geeksforgeeks.org/deep-q-learning/

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., Van Den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359. https://doi.org/10.1038/nature24270

Zhao, M., & Marquez, A. G. (2013). Understanding humans’ strategies in maze solving. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1307.5713