7-3 Project Two: Design Defense

Paloma Rodriguez

CS-370 : Current/Emerging Trends in CS

October 9th, 2023

The purpose of this assignment is coming up with a strategy to navigate through a pirate theme maze. We used a python Deep Q Network to work with two other classes. The pirate character can only move in four directions and is rewarded for successfully reaching the treasure and penalized for hitting obstacles, revisiting the same location, or attempting an invalid move. Through reinforcement learning, the pirate is trained to pick the most optimal action for each state which is done through exploration and exploitation.

When talking about exploration with reinforcement learning this revolves around the idea of selecting actions that the agent has not yet attempted or it believes will provide new insights to enhance its long-term performance. Its purpose is to gain new knowledge and information that can lead to better decision-making in the future.

Exploitation involves the strategy of choosing actions that the agent believes will have the highest immediate reward based on its existing knowledge. Exploitation focuses on maximizing short-term performance by making decisions grounded in the current understanding it has.

Now to gain a happy medium between exploration and exploitation we can implement a learning rate decay mechanism. This decay rate gradually shifts from random exploration to informed exploitation as it gains experience

During the early stages of learning, priority is given to exploration to allow the agent to explore different paths and understand its environment. As the agent accumulates more experience and knowledge, the balance shifts towards exploitation to leverage that knowledge for improved performance.

Reinforcement learning assists in determining the path to the goal (the treasure) for the agent (the pirate). The agent learns through its actions, forming states to actions. During pathfinding, states will relate to various locations on the map, and actions represent the movements the agent can make to navigate the maze. The agent is going to maximize the total expected reward over time by adjusting its actions based on exploration and exploitation strategies.

References:

Hadsell, R., Rao, D., Rusu, A. A., & Pascanu, R. (2020, December). Embracing Change: Continual Learning in Deep Neural Networks. *Trends in Cognitive Sciences*, *24*(12), 1028–1040. https://doi.org/10.1016/j.tics.2020.09.004

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

Lamba, A. (2018, August 27). *A brief introduction to reinforcement learning - We’ve moved to freeCodeCamp.org/news - Medium*. Medium. https://medium.com/free-code-camp/a-brief-introduction-to-reinforcement-learning-7799af5840db