Mitchell Fitzsimmons

4/18/2024

CS 370

Project 2

There are noticeable similarities and differences in the approach humans and machines make in problem solving, specifically in the context of the maze presented. Humans take a problem and break down the different avenues they could go about solving it and pick what they think is the best solution based on what they intuitively think, which can be influenced from past decisions. In context with the maze, they would pick a path that they think would most likely be correct, and if it is not then would go back and continually apply this process until they successfully reached the end of the maze. For an artificial intelligence system rather than thinking and assuming a correct path, it would choose paths at random until it found success then output whatever the successful path to the end of the maze would be. The AI system learns from the mistakes it makes rather than weighing possible consequences of actions.

These systems are different in the approaches mentioned above but do share fundamental similarities. In example, both a computer and a human before deciding needs input, the presented problem, and ways to solve the problem. While humans can immediately make choices based on possible consequences, the AI system will do actions at random, and then learn of the consequences of said actions that were selected at random. Another similarity is the end result. The end result of both of these systems is the maze being completed, and both systems learning and having a better understanding of how to complete the maze, and possibly other mazes as well.

Exploration and Exploitation have key differences, which can be outlined in the context of this problem. Exploration can be defined as solving problems through searching and finding solutions through trial and error. This could be seen as the human way of taking a problem step by step and measuring the consequences step by step. Exploitation can be defined as taking every possible route to solve a problem and choosing the most optimal solution. While ideally, an AI system should be able to use a combination of both these ideas, usually the brute force method of exploitation is most commonly used. In reinforcement learning, the idea of trial-and-error exploration alongside the exhaustive exploitation can help this system learn incredibly fast. Having the system learn what paths are correct and incorrect can help it make judgements in the future to successfully complete the maze. This can be as simple as best directional movements (movements in the general direction of the end of the maze can be identified as most intelligent), or as complex as the idea of constantly moving right in a maze will eventually give you a correct solution.

Deep Q-learning was implemented for the pirate agent to best finish the maze and solve the problem presented to it. In the context of the provided .ipnyb file provided this can be broken down into clear steps. First, implementing the maze and having a method to visualize the maze can help understand what is going on, alongside having the Keras library import for the functions that are called throughout the program. Next the system will need to train through the process of randomly selecting paths until it either fails or succeeds. From here the system learns based on the built-in reward system informing the AI whether that run was a success. This allows the AI to build from there and get better at completing the maze resulting in a better ‘reward’. After this testing agent is done, the program should be able to output what is the best path based on the ‘reward’ system. This is how the Deep Q-learning system works at 10,000-foot overview.

Works Cited:

Amine, A. (2020, December 19). *Q-Learning Algorithm: From Explanation to Implementation*. Medium. https://towardsdatascience.com/q-learning-algorithm-from-explanation-to-implementation-cdbeda2ea187

‌ *Q-Learning in Python - GeeksforGeeks*. (2019, February 6). GeeksforGeeks. <https://www.geeksforgeeks.org/q-learning-in-python/>

Rocca, J. (2021, May 23). *The exploration-exploitation trade-off: intuitions and strategies*. Medium. https://towardsdatascience.com/the-exploration-exploitation-dilemma-f5622fbe1e82

‌

‌