Travis S Waggoner

Southern New Hampshire University

CS-370

10/10/2022

Project Two

The steps that a human would take in order to solve the maze would be simple. This would look like blindly picking a direction to start and working their way through the maze one direction at a time. This would result in them finding what directions that they are able to go and not able to go through trial and error. There would not be any specific methodology for solving this problem through the average person. As the human can successfully find the correct spot in the correct number of moves, they will be able to find the best solution for the problem.

The intelligent agent solves the problem be picking action and moving in that direction. As it progresses through the maze it will find what actions that are invalid and valid actions through trial and error. This results in the agent being able to move in a methodical manner resulting in an accurate grouping or solutions. Once the agent has found the solution it is able to find the best answer to the problem. This results in a solution that takes the least number of moves resulting in the best reward to the agent.

These two methods are very similar because there is a amount of information that is not known to there agent. This means that there is a amount of “guess” work needed to reach the final goal. There is also similarity to the moves that both agents can make within the maze while this is not a result of the agent themselves this does create similar circumstances. The differences between the agents are how they find the answer and how methodical finding the correct answer is. The agent will result in a more accurate answer it a theoretical shorter amount of time.

Exploration is when the program is attempting to find new methods of solving the problem for a greater reward. Exploitation is when the program takes advantage of known data to create a better result for the best reward. In this case Exploitation would be able to solve the problem faster because the answer at the beginning would be unknown with few known variables. The exploitation method would be great at finding the best answer for the problem once the problem has been solved or is close to being solved.

Reinforcement learning would help because you could use the hot and cold method for the AI to learn the problem better. This could in turn look something like when the pirate moves it is rewarded and if it moves closer to the solution it gets a better reward verse if it moves farther away from the reward it gets a smaller reward. If the AI does not move or moves to a spot that it has already been on it does not get a reward. This makes the pirate move towards the reward faster but does not make it so easy as to not learn the solution.

Deep learning is implemented in the game by implementing how the agent will interact with the puzzle to learn the correct route. This can be seen on how the actions are designed to be valid or not and how the reward is given to the AI. This also provided the reward to the AI on what the answer will be within the neural network. This information is than saved as experience and then provided as the input data.

Work Cited

Beysolow, I. T. (2019). *Applied reinforcement learning with python : With openai gym, tensorflow, and keras*. Apress L. P..

Gulli, A., & Pal, S. (2017). *Deep learning with keras : Get to grips with the basics of keras to implement fast and efficient deep-learning models*. Packt Publishing, Limited.

Lamba, A. (2018, September 3). *An introduction to Q-learning: Reinforcement learning*. Medium. Retrieved October 16, 2022, from https://medium.com/free-code-camp/an-introduction-to-q-learning-reinforcement-learning-14ac0b4493cc

Matthew Botvinick, Sam Ritter, Jane X. Wang, Zeb Kurth-Nelson, Charles Blundell, Demis Hassabis, Reinforcement Learning, Fast and Slow, Trends in Cognitive Sciences,Volume 23, Issue 5, 2019, Pages 408-422, ISSN 1364-6613, <https://doi.org/10.1016/j.tics.2019.02.006>. (https://www.sciencedirect.com/science/article/pii/S1364661319300610)