Merlin Martinez

08/15/2024

Module 7.3

Project Two

Some of the steps that a human being would take to solve a maze would be trying to look around and try to, if possible, to climb to be able to see the maze from the top. The second thing would be going in different directions and if leads to nothing go back to the starting point until the end of the mazed is reached and go on until finding the fastest route.

The IA doesn’t think like human beings so they would take different steps because they will be exploring the mazed by trials and errors. The IA is going to take random routes so they can find the quickest way.

Retrieve data from the maze like the start and end

IA goes through the algorithm

After the IA goes thought different paths multiple times, they will be able to find the fastest route.

Both humans and IA have to go through the maze multiple times until the fastest route is found however the IA different from a human being would use number array as well different from humans if it has the rewards penalty system it will try its best to try and obtain the highest reward the IA has not knowledge until later on and its because it learns from trials and errors, humans in other hands would use their instinct as well visual queues to try and not go through the wrong direction most of the time humans would try to find linear paths we all do whenever we are lost.

The difference between the two are the following exploitation is more about obtaining a bigger and immediate rewards different from exploration which is more about finding solutions by searching, testing and experimenting and this means that the rewards are even bigger. In conclusion both are needed and necessary. I think the ideal proportion of exploitation and exploration for this pathfinding problem would be doing penalty moves every time the agent will try and make a move like getting out of the path or going through the walls etc.

Reinforcement learning can help to determine the path to the goal by the agent by doing trial error and increasing its reward and penalty so it can guide the agents’ actions, this will ensure the agent doesn’t do any wrong or random moves.

The way I implement deep Q-learning using neural networks for this game is by using the agent neural network to store Q-values as well creating a maze environment and so they agent neural network as well I randomly selected a free cell as the starting position, then IA could see the game environment to do valid actions, the IA uses a combination of exploration, exploitation strategy. After the new state has been observed we can use the Q-value of the action it chooses and then it gets updated in the Bellman equation. After that it gets repeated and updated multiple times until the desired outcome.

Resources

Karagiannakos, S. (2018, November 17). The idea behind Actor-Critics and how A2C and A3C improve them | AI Summer.AI Summer. <https://theaisummer.com/Actor_critics/>

Gautam, A. (2018). Introduction to Reinforcement Learning — Part 1 - Adesh Gautam - Medium.Medium. [https://adeshg7.medium.com/introduction-to-reinforcement-learning- part-1-dbfd19c28a30](https://adeshg7.medium.com/introduction-to-reinforcement-learning-%20part-1-dbfd19c28a30)

Loeber, P. (2022, February 10). Reinforcement learning with (deep) Q-learning explained.

News, Tutorials, AI Research. <https://www.assemblyai.com/blog/reinforcement-learning-with-deep-q-learning-explained/#:~:text=Deep%20Q%20Learning%20uses%20the,action%20based%20on%20that%20state>.