CS – 370

Shane Griffith

6/15/2022

Week 7 final Project

As a human attempting to traverse this maze, I would use the sensory option available to do the best I could to find the treasure. Most importantly would be sight. After being placed in the maze, I would be able to see which direction I could move to accurately move through the maze to find the treasure. The intelligent agent in this exercise was able to find the treasure by getting either rewarded or punished based on the movement that it took. For example, if the agent took a step that was either into a wall or further away from the treasure, then it would get penalized for that by decreasing the agents score. On the other hand, if the agent moved in a direction that was closer to the treasure, then the agent was rewarded by increasing the score. The agent was ultimately able to traverse the maze by learning and only making movements that would put it closer to the treasure and not further away.

Exploitation in this sense would be the agent simply moving around the maze and learning where to move based on the result of the movement itself. For instance, if the agent moved and the score increase, the agent then knows it made a move towards its end goal and can learn form there in which direction the treasure is. Let’s assume there were 2 agents in this scenario. If agent A were to traverse the maze and find the treasure and agent B heard where agent A found the treasure, then Agent B could get greedy and go and search the same place the treasure was found by agent A to see if it was there again. This would be an example of exploitation. To use all of the information already available to make a decision opposed to a strategy of exploration that would be much more based on the result of the previous action.

Reinforcement learning would an excellent way to resolve the pathfinding problem. As explained in the previous paragraph, the agent could be “reinforced” to make its next decision based on the result of the previous steps. If the agent decides to move in a direction that results in points towards its high score, then the agent knows it is moving in the right directions. However, on the other hand if the movement the agent decides to make yields in negative or no points, the agent then knows it is moving further away from the treasure and will have to make an adjustment to complete the maze and find the treasure.

Q-Learning can be implemented in this exercise by generating a matrix that would be expressed as the maze for the agent to move through. As explained in the provided material, an 80% grey square would be left behind the agent as it moved towards the treasure. The 100% black square would will represent walls or obstacles the agent would have to traverse around and the 100% white squared would remain untouched by the agent.