CISC453 Final Project Proposal - Battleship

Group 3

Jacqueline Heaton   20028278
Katy Scott   20009328
Kenya Moore   20020515
Julia McPolin   20020258
Rachel Theriault   20005337

Description: In this project we will design two agents that use different learning techniques to play the game Battleship. The board will be a 5x5 grid and each player will have three ships, two of size 2 (submarine) and one of size 3 (cruiser). Ships will be placed randomly for both players. The policy will be updated as the agents learn how to hit the other player's ships. Each agent will know the size of the ships, if they hit or miss on a turn, and if they sink a battleship. We will have a stochastic base policy that will always randomly select an action, and the two policies will be trained against this. The two trained agents will then play against each other and we will compare the results. The only action available to the agent is selecting a square to fire a shot at.

Rewards: The agents receive a negative reward for a miss and a positive reward for sinking a ship.

Objective: The agent will learn to sink all of their opponent's battleships in the least number of moves.

Learning technique: We will be using Monte Carlo and Q-learning algorithms. They will be trained against a stochastic agent.

Implementation: The implementation will be done in Python 3. We will use Tkinter for our UI.

Deliverables: The project will generate the following deliverables:

1. A graphical representation of each player's board. Ships will be represented as occupied cells on the board. Missed shots will be represented by a white x and hits will be represented by a red circle.

2. A learning environment to generate each agent's experience. Each learning agent will be trained against our base policy agent. At certain episodes the learning agent's will be pitted against each other to see how well their policies work.

3. Data showing the comparison between the learning techniques at different numbers of training iterations and plots of the data.

4. A report describing the algorithms used, and results shown as graphs demonstrating the learning of the task.