Student ID: _____
Name: _____
Mark: _____

COMP6247 Lab - Weeks 9 and 10

Adversarial learning

Due 09/05/2019

Consider the 2-player capture the flag game from labs 7-8 again. Just a reminder: the game consists of 5 sites and 2 flags. At each round, Player 1 (the hider) chooses 2 sites to hide her flags. Note that each site can contain at most 1 flag. Player 2 then chooses 2 sites to check whether the flags are hidden there. If the seeker finds both flags, she gets 2 points, if she finds only 1 flag, then her reward is 1, and 0 otherwise. The goal of the seeker is thus to maximise her total points over time, while the hider aims to minimise the total reward of her opponent.

1. [*10 marks*] Consider the following asymmetric game setting: Player 1 can see what Player 2 has chosen each round, while Player 2 can only see the places of the flags if they were at the chosen sites of Player 2. Please choose an appropriate learning algorithm for both players and implement them against each other. Play the game for 500 rounds, and repeat this for 100 times. Plot the average reward of the seeker (i.e., the total reward divided by the number of plays) over number of plays. Can you observe convergence of this value at all?

2. [*10 marks*] Now suppose that the hider can lie about the position of the true flags during the game. For example, if the seeker chooses sites 1 and 2, and the flags were in fact on sites 3 and 4, the hider can still say that the seeker found the flags and received a reward of 2 (the hider can also lie about 1 site, not 2). Basically, when the hider lies, she can give any arbitrary information. To restrict the hider's power, we assume that she can lie up to B times during a game. Can you come up with a smart faking strategy for the hider such that she can manipulate the seeker to have a low performance in the end (after revealing the fake flags)? Once you have the faking strategy, please play the game with $B = 10, 50, 100$ for 500 rounds, and repeat this for 100 times. Plot the average reward of the seeker (i.e., the total reward divided by the number of plays) over number of plays, both before and after the reveal for all three cases ($B = 10, 50, 100$). Also, please plot the difference between the true total reward (without the fake flags) and the total reward the seeker thought se received at the end of the game.