Week 4 report

Understand the Deep Reinforcement Learning

What I have done:

* Read through the literature about deep RL applying in battleship game.
* Read through the code it provides.
* Rewrite the code to class-style, so that it could be modified easier for later use.
* Went through the tutorial of Tensorflow

What I have learnt

* We typically use the reinforcement learning technique in a program to play games effectively.
* In this example, past game data is used to train RL.
* Principle of the RL in this case is that
  + let the program to recognise good strategies in the data.
  + Make the program learns new strategies fast when runs the game.
* Policy Network and rewards function are 2 main elements of deep RL.
* Policy Network is a neural network that given an action α, it maps the state values *s* to probabilities.
* Rewards function is used to score the outcomes of past games.
* Stochastic gradient descent is used here to train the program to optimize the rewards while playing the game.
* The code provided has simplified the game by making as 1d map with size of 10 and 1 ship with size of 3.
* Tensorflow is used in the code for easier coding RL process.
* Softmax regression is a simple model to assign the possibilities to an object being many different things.
* Cross-entropy is the score of the data from its optimal model runs with another optimal model. It describes how different 2 models are.
* A more complex model called Multilayer Convolutional Network can be more accurate.

What I plan to do next week

* Modify the code to make it 2d area space and add more ships to makes it looks more like real game. (which has much more search space than the one in the example) To experiment more complicate game setting.
* The AI in the example code takes the board state as input which is not tally the description of ‘high dimension level sensor input’, I will try to code the AI so that it actually takes the pixels as input.
* Find out more examples of the deep RL applying in other games.