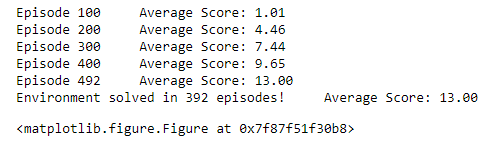
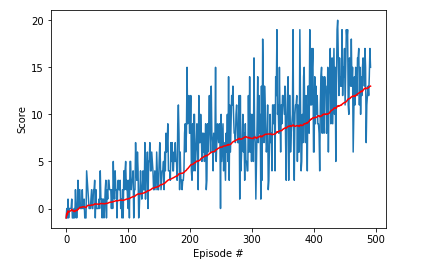
P1 Navigation Project Report

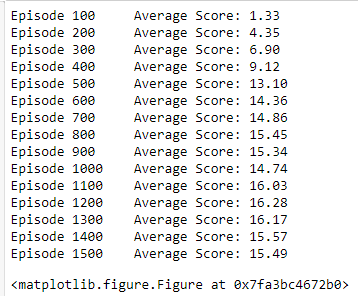
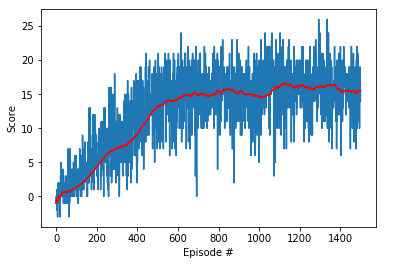
Learning Algorithm and Plot of Rewards:

* The learning algorithm is DQN, a value based RL learning algorithm.
* The DQN agent model is a neural network which includes following layers
  + 3 Linear (Dense) layers
  + 1st and 2nd linear layers are followed by relu activation layer
  + 3rd linear layer is output layer which outputs the value for each stage-action pair
* Here are the hyperparameters used in 1st training experiment and experiment results.
  + eps start=1.0, eps end=0.01, eps decay=0.995
  + BATCH\_SIZE = 64, GAMMA = 0.99, TAU = 5e-4, LR = 0.0005, UPDATE\_EVERY = 4

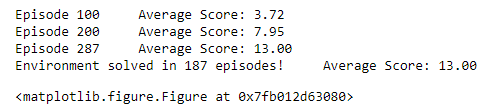


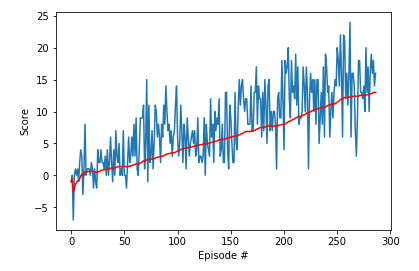


* Change the threshold from 13.0 to 20.0

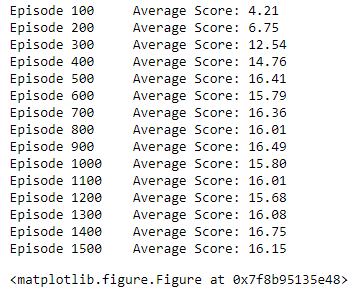
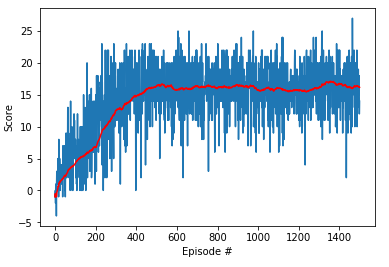


* Here are the hyperparameters used in 2nd training experiment and experiment results.
  + eps start=1.0, eps end=0.01, eps decay=0.5
  + BATCH\_SIZE = 64, GAMMA = 0.99, TAU = 5e-4, LR = 0.0005, UPDATE\_EVERY = 4





* Increase the stop threshold from 13.0 to 20.0. The training saturates at average reward 15.0



* Conclusion:
  + It seems that reducing eps decay, which accelerates decaying rate of exploration, will help the agent converge faster.
  + I also tested different learning rate and the result indicates that by increasing learning rate from 0.001 to 0.0005, the final convergence of the reward will increase from 15.0 to 16.0

Learning Algorithm and Plot of Rewards:

* Change DQN to DDQN
* Apply Prioritized Experience Replay
* Try Dueling DQN
* Apply Actor-Critic learning algorithm
* Increase layers of original DQN neural networks