Choo Han Ye 1002439

AI Coding Homework Week 9

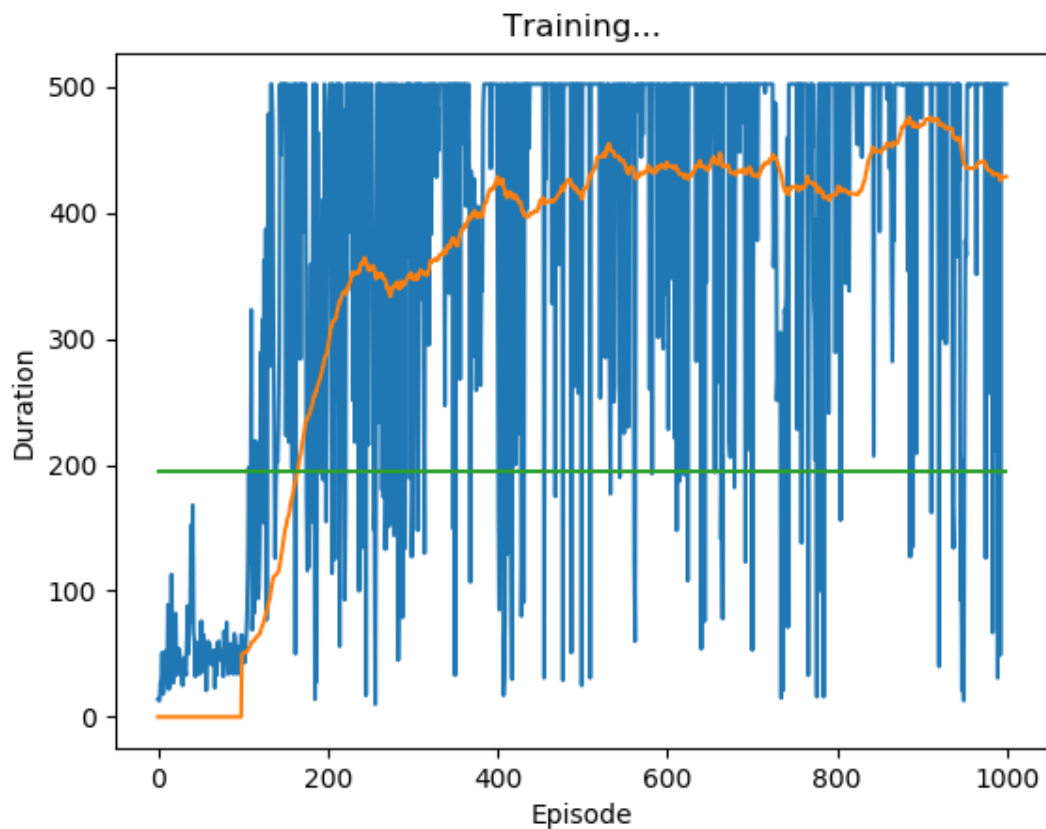Task 1

Below are the parameters used:

GAMMA = 0.999

EPS = 0.25

LEARNING_RATE = 0.005

WEIGHT_DECAY = 0.000005

Neural Net: 1 hidden layer with 128

Graph:

Maximum averaged duration achieved is 475.91. This is much better than last week's cartpole done. Target duration achieved at around 150th episode.

Task 2

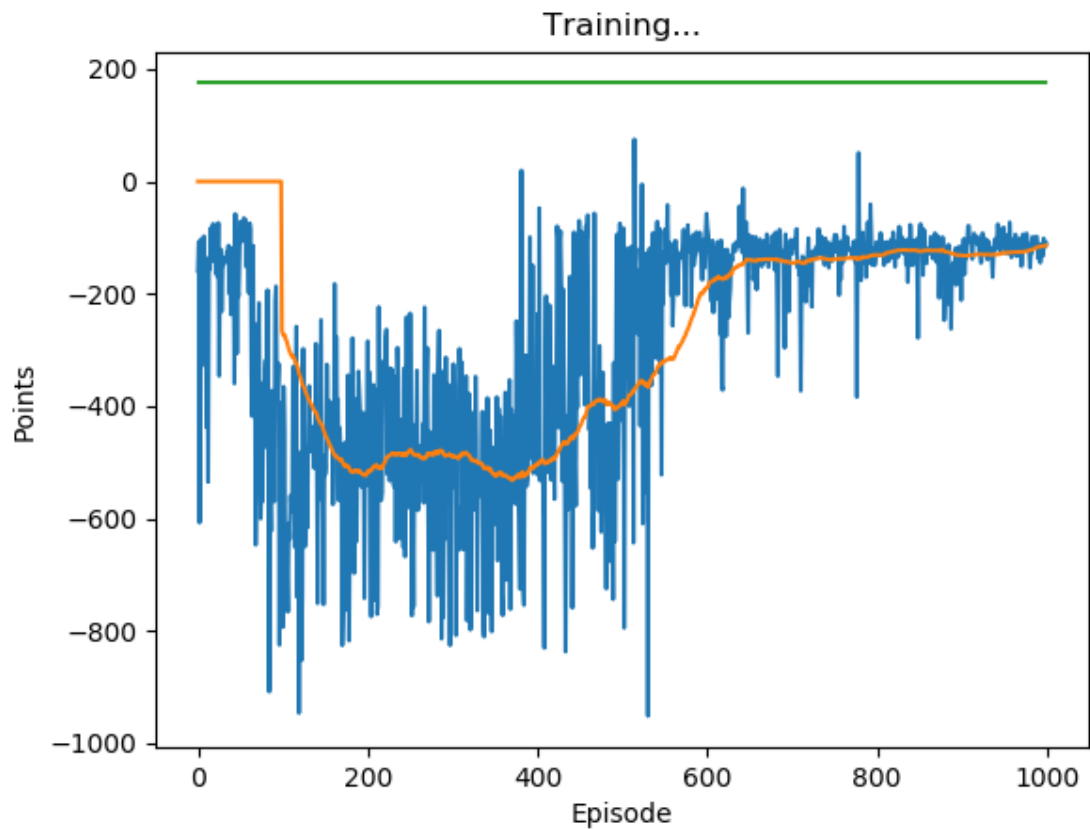Below are the parameters used:

GAMMA = 0.999

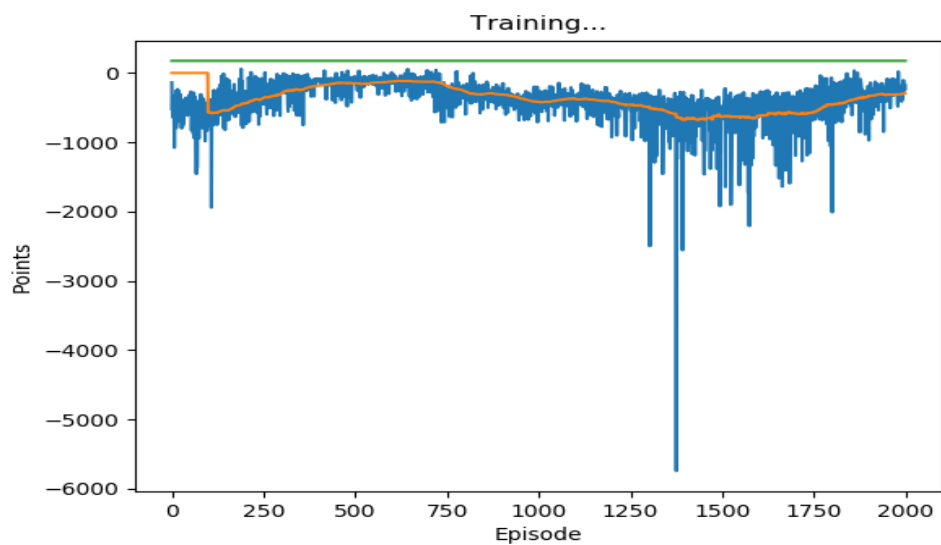EPS = 0.25

Choo Han Ye 1002439

LEARNING_RATE = 0.005

WEIGHT_DECAY = 0.000005

Neural Net: 1 hidden layer with 128

Graph:

Within 1000 episodes, it was nowhere near the target. So I tried again with 2000 episodes:

Choo Han Ye 1002439

Neither 1000 nor 2000 episodes could achieve above 0 average points....

Task 3

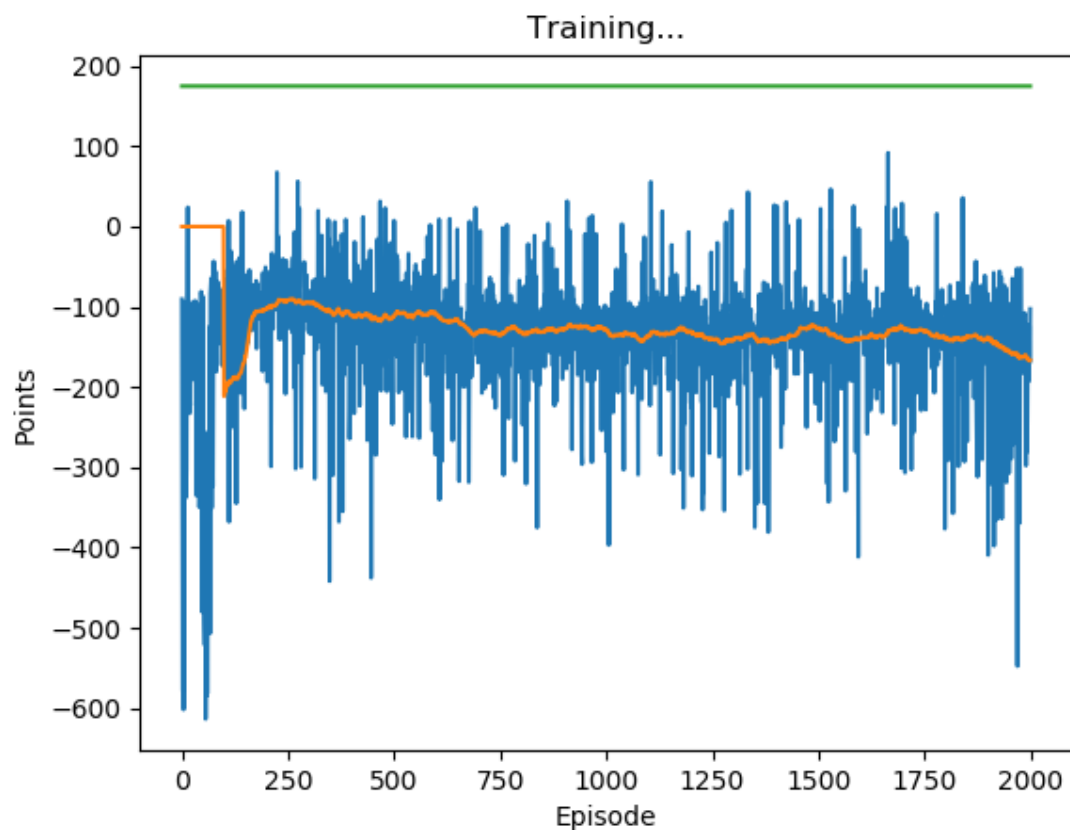Below are the parameters used:

GAMMA = 0.999

EPS = 0.25

LEARNING_RATE = 0.005

WEIGHT_DECAY = 0.000005

Neural Net: 1 hidden layer with 128

2000 Episodes

Graph:

As we can see, even though performing RL with and without baseline both does not achieve above 0 averaged points, performing RL without baseline achieves a worse result (Task 2 approaches 0 while Task 3 approaches -200).