Deep Reinforcement learning Project 1 "Navigation"

Dongsu Du

1 Deep Q-Network Architecture

I decide to use the layers as 1st hidden layer 128, 2nd hidden layer 64, 3rd hidden layer 32. These choices can give me pretty good results without too much calculation time. Relu function is selected as the activation function between layers. I also tested the 2 layers with 128 and 64 hidden layers, 2 layers with 256/128/64 hidden layers. The result shows that the total episodes needed is

Hidden layers	128/64/32	128/64	128	256/128/64
Episodes	358	317	432	421

Table 1. The

2 DQN Agent algorithm

This algorithm I used is just the deep Q learning algorithm in the lecture (without double DQN or dual DQN. The agent already can find the good solution within 150 iterations. The constraint I setup for stop training is the average score >15. For each episode, the maximum steps = 1000.

2.1 Testing on value of eps

The choice of eps decay speed is important in the training. In table 1, I summarized the total episodes to achieve average score > 15 for different eps selection.

Eps decay	0.99	0.98	0.97	0.95	0.9
Episodes	391	386	358	334	323

Table 2. The total episode with different eps decay rate.

3 Result

All the training only use CPU. The final result is shown below (eps decay = 0.9, Hidden Layer = 128/64/32)