Project 1: Navigation

Results

My trained agent was able to successfully meet specifications by performing 100 sessions with average scores of ~15 points, exceeding the target of 13 points.

Solution

The agent was implemented as a DQN with 2 fully connected hidden layers with 64 nodes each.

Actions were chosen in an epsilon-greedy fashion starting with an epsilon value of 100% which was gradually annealed to 1% using a epsilon decay of 99.5%.

Other parameters include:

BUFFER_SIZE = int(1e5) # replay buffer size
BATCH_SIZE = 64 # minibatch size
GAMMA = 0.99 # discount factor
TAU = 1e-3 # for soft update of target parameters
LR = 5e-4 # learning rate
UPDATE_EVERY = 4 # how often to update the network

Training was performed for 1500 episodes. After 1000 epsidoes performance improvements were observed to stall.

Further Work

The above hyperparameters were not extensively explored, and could certainly be optimized.

The network design is also not optimized, and the number of hidden units is certainly not optimal.

Finally, as the implemented algorithm is a basic DQN design, many standard enhancements should prove beneficial. The implementation could be expanded to a full blown Rainbow algorithm.