Navigation

Learning Argorithm   DDQN(Double Deep Q Network)

parameter	replay buffer size	100000
	batch size	64
	discount factor	0.99
	soft update of target parameters	0.001
	learning rate	0.0005
	how often to update the network	4
	maximum number of training episodes	2000
	maximum number of timesteps per episodes	1000
	starting value of epsilon, for epsilon-greedy action selection	1
	minimum value of epsilon	0.01
	multiplicative factor (per episode) for decreasing epsilon	0.995
neural network	state size	37
	action size	4
	number of nodes in first hidden layer	64
	number of nodes in second hidden layer	64

Plot of Rewards

Reward values during training

Reward value when using learning results
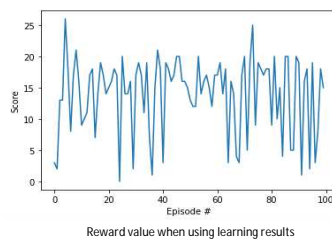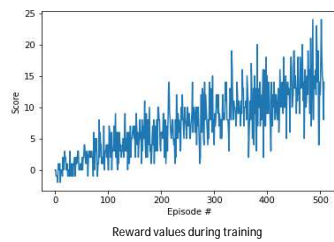
Ideas for Future Work

Dueling Network and Prioritized Experience Replay can be used as improvement measures
to speed up learning.