■ Navigation

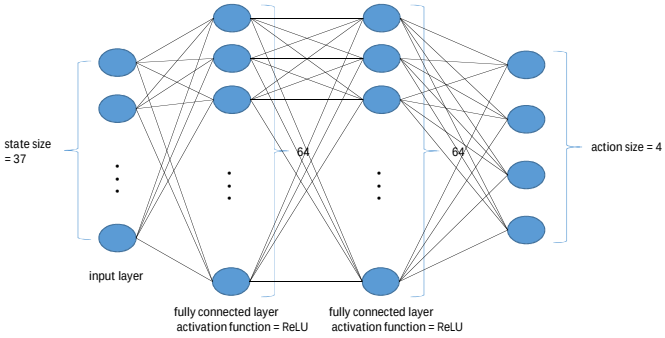
• Learning Argorithm　　DDQN(Double Deep Q Network)

Q-learning is a type of value based method, in which the value of an action taken in a certain state is managed in a table called the Q table, and the Q value is updated each time the action is taken. DQN is a method that applies a neural network to Q-learning and learns efficiently even when there are many states and actions. In DQN, if the target value is set too large when calculating the TD error, there is a problem that the previous state is overestimated.Therefore, Double DQN is designed to stabilize the calculation of TD error by training by mixing two Q functions.

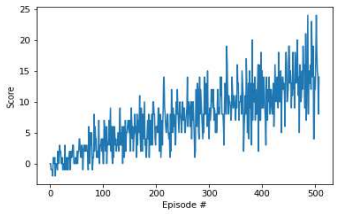
• Hyperparameters

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

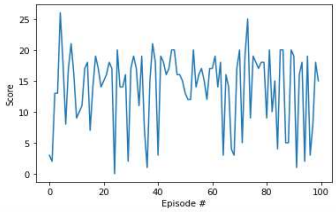
• **Model Architecture**



• 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.