

Report for Assignment 3

z5099208 Jiazheng Huang

z5085901 Jingjie Jin

1. In this assignment, we firstly implemented **basic DQN network with random batching and experience replay**. Then after parameter tuning, we found the DQN is not stable enough and the reward at 100 steps is not high which is around 60.
2. Then we started to optimize it by implementing **dueling DQN** when defining the network. As a result, we found that this one is learning faster but it was still unstable.
3. In this step, we thought the unstable situations might be caused by overestimating. Then we implemented the **double DQN** to avoid it. In the double DQN, we also use **fixed parameters** of target network and update it each 50 steps to improve the stability of the model. Finally, we get a kind of stable model.