

Correction Notes - Assignment 4

Group: Manzil, Cüneyt and Dhananjay

Task 1

$TD(0)$ is correctly implemented. The results are also plausible. Good job on this task.

Points: 10/10

Task 2

You misunderstood the initialization of Q and do not use your results from task 1 here (-1 pt). Apart from that, your exploratory actions follow the policy from task 1, which was not intended, but is actually probably better here. Q -learning is implemented according to the lecture. Your results are not very good, but this was very difficult in this task. Overall good solutions on this whole sheet.

Points: 9/10

Total Points: 19/20 Points