2) Once you have your robot working, measure its learning performance as follows:

a)  Draw a graph of a parameter that reflects a measure of progress of learning and comment on the convergence of learning of your robot.

b)  Using your robot, show a graph comparing the performance of your robot using on-policy learning vs off-policy learning.

c)  Implement a version of your robot that assumes only terminal rewards and show & compare its behaviour with one having intermediate rewards.

3) This part is about exploration. While training via RL, the next move is selected randomly with probability  and greedily with probability 1 −

a) Compare training performance using different values of  including no exploration at all. Provide graphs of the measured performance of your tank vs .

As for part 1, your submission should be a brief document clearly showing the graphs requested about. Please number your graphs as above and also include in your report an appendix section containing your source code