Addendum

In addition to the RL assignment:
 - The weights on the connections are to be understood as "reward".
 - the objective is to maximise the reward
 - consider the "gamma" parameter (discount factor) as 1