

REINFORCEMENT LEARNING Exercise 2



1 Markov Decision Processes

Assume the following problem: there are 5 parking spaces and you start at parking space 5. In each step, you can either try to park or drive on. A parking space is free with probability p . If a parking space i was occupied or you drove on, you move to the next parking space $i - 1$. You want to be as close to your home – which is at parking space 1 – as possible. However, you want to avoid to reach the end of parking spaces without parking successfully.

- (a) Formalize the above problem as an MDP.
- (b) Draw the transition graph.
- (c) Do we have to discount? Explain your answer.

2 Markov Property

Assume a biased slot machine in a casino. Each round, the player can win 1\$. However, whenever the outcome of the last two rounds is larger than 1\$, the machine lowers the probability of winning. Is the Markov property fulfilled?

3 Optimal Value Function

- (a) Disprove by counter example: for any MDP with optimal value function v_* , the optimal deterministic policy π_* is unique.
- (b) Disprove by counter example: the optimal value function v_* for state s_t at time step t is always larger than for state s_{t+1} under the optimal deterministic policy π_* , i.e. $v_*(s_t) > v_*(s_{t+1})$.

4 Experiences

Make a post in thread *Week 02: Markov Decision Processes* in the forum¹, where you provide a brief summary of your experience with this exercise and the corresponding lecture.

¹https://ilias.uni-freiburg.de/goto.php?target=frm_1380029&client_id=unifreiburg