MSc in Computer Science and Engineering

Learning and Decision Making 2017-2018

Homework 2. Markov decision problems

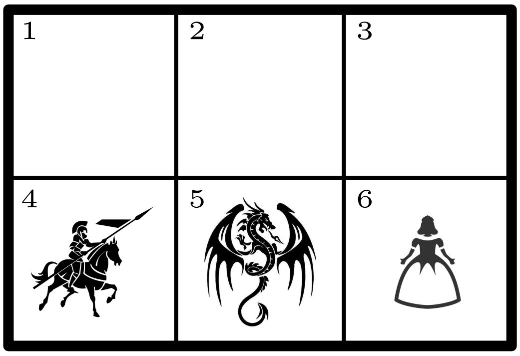

Figure 1: Grid environment where a knight must save a princess and avoid the dragon. The princess and the dragon do not move.

Consider a knight seeking to save a captured princess in a 2×3 grid world (see Fig. 1). The princess is held in position 6 of the grid, while a dragon guards the passage in position 5 of the grid.

At each step, the knight may move in any of the four directions—up, down, left and right. The movement succeeds with a 0.6 probability and fails with a 0.4 probability. When the movement fails, the knight may stay in the same cell or move to one of the immediately adjacent cells (if there is one) with equal probability.[1] See Fig. 2 for some examples.

The goal of the knight is to save (reach) the princess and avoid the dragon. In this homework, you will model the decision of the knight as a Markov decision problem (MDP).

Exercise 1.

(a) Identify the state space, $\mathcal{X}$, and the action space, $\mathcal{A}$, for the MDP.

[1]When that implies going beyond the limits of the grid, the corresponding probability adds to the probability of staying in the same cell.

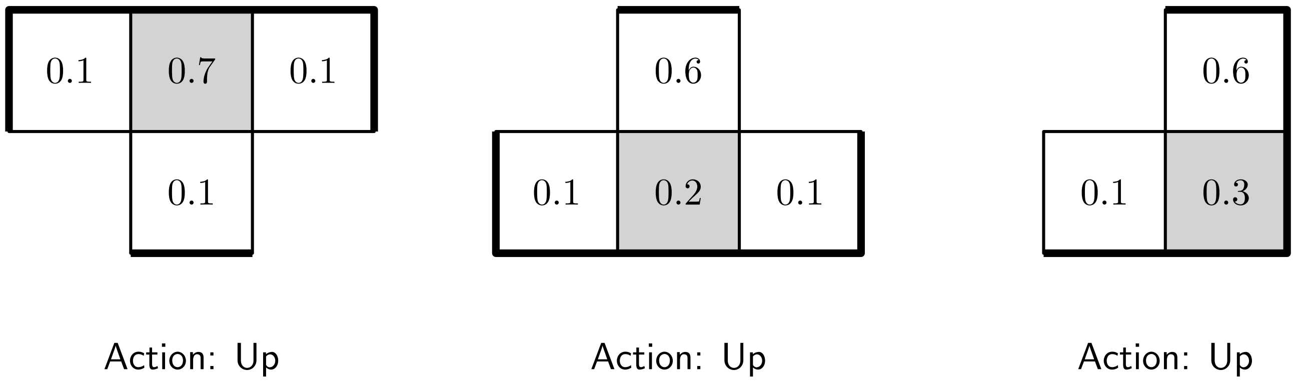

Figure 2: Examples of how transition probabilities change depending on the knight's position in the grid.

(b) Write down the transition probabilities and a (possible) cost function for the MDP. Make sure that the cost function is as simple as possible and verifies $c(x, a) \in [0, 1]$ for all states $x \in \mathcal{X}$ and actions $a \in \mathcal{A}$. Note, in particular, that the cost should depend only on the knight *standing* in the same cell as the princess (success) or the dragon (penalty).

(c) Compute the cost-to-go function associated with the policy in which the knight always goes up, using a discount $\gamma = 0.9$. You can use any software of your liking for the harder computations, but should indicate all other computations.