MSc in Computer Science and Engineering

Planning, Learning and Decision Making

Homework 2. Markov decision problems

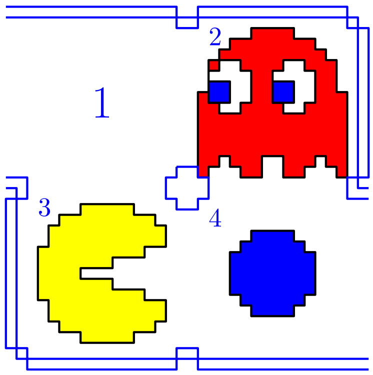

Figure 1: Pacman and ghost in a 2×2 grid.

Consider the snapshot depicted in Fig. 1, representing Pacman and a ghost from the Pacman game. In cell 4 you can also see a blue "pellet", that gives Pacman special powers. Unlike the first homework, in this case assume that the ghost *does not move*.

In this homework, you will describe the game of Pacman using a Markov decision problem. In this case, the decision-maker (the player) controls the movement of the Pacman character. As a player, you have available 4 actions: "Up", "Down", "Left", and "Right", each of which moves the Pacman character one step in the corresponding direction, if an adjacent cell exists in that direction. Otherwise, Pacman remains in the same place.

Pacman can be in any of the 4 numbered cells; the cell in the top left corner (cell 1) is adjacent, to the left, to the cell in the lower right corner (cell 4). In other words, if the Pacman "moves left" at cell 1, it will end up in cell 4, and vice-versa.

If Pacman lies in the same cell as the ghost (cell 2), the player looses the game (i.e., the game should transition to a "Defeat" state). However, if Pacman "eats" the blue pellet, it gains the ability to "eat" the ghost. In this case, if Pacman lies in the same cell as the ghost, it "eats" the ghost and wins the game (i.e., the game should transition to a "Victory" state). Assume that Pacman can never be in cell 4 without "eating" the pellet.

Exercise 1.

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

(b) Write down the transition probability matrix for the action "Left" and a (possible) cost function for the MDP. Make sure that:

- Both victory and defeat are *absorbing states.*
- 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 $a \in \mathcal{A}$.
- The cost should depend only on *where Pacman is.*
- Both absorbing states (victory and defeat) should have a cost of 0.

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