MSc in Computer Science and Engineering

Planning, Learning and Decision Making

Homework 3. Partially observable Markov decision problems

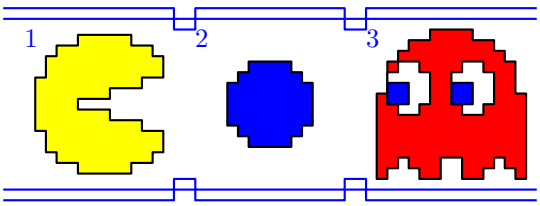

Figure 1: Pacman and ghost in a 1×3 grid.

Consider the snapshot depicted in Fig. 1, representing Pacman and a ghost from the Pacman game. In cell 2 you can also see a blue "pellet", that gives Pacman special powers.

In this homework, you will describe the game of Pacman using a partially observable Markov decision problem. In this case, the decision-maker (the player) controls the movement of the Pacman character. As a player, you have available 2 actions: "Left" and "Right", each of which moves the Pacman character one step in the corresponding direction.

In this version of the game, Pacman can be in any of the three numbered cells. The leftmost cell (cell 1) is adjacent, to the left, to the rightmost cell (cell 3) and vice-versa. The ghost moves back and forth between cells 2 and 3 in a deterministic manner.

If Pacman lies in the same cell as the ghost, 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 2 without "eating" the pellet.

Finally, Pacman is unable to see the ghost unless if it stands in the same position as the ghost (however, it does know its own position and whether it ate the pellet or not).

Exercise 1.

(a) Identify the state space, $\mathcal{X}$, and the action space, $\mathcal{A}$, and the observation space $\mathcal{Z}$ for the POMDP. As in Homework 2, you should treat the "Victory" and "Defeat" states as absorbing states.

(b) Write down the transition and observation probability matrices for the action "Left".

(c) Suppose that, at time step $t = 0$, Pacman is in cell 1. It has not eaten the pellet and does now know where the ghost is. It takes the action "Left" and moves, accordingly, to cell 3 at time step 1, but does not observe the ghost. Indicate the initial belief state of the agent and the updated belief at time step $t = 1$.