Homework 2. Markov Decision Problems

Grupo 27
97155 João Aparício
97040 Rostislav Andreev

Exercise 1

(a) Each state is defined by the number of the cell on which the taxi is currently on and by whether or not the passanger is in the car.

$$\mathcal{X} = \{(1,P), (2,P), (2,\neg P), (3,P), (3,\neg P), (4,P), (4,\neg P)\}$$

Each action represents a direction in which the taxi may move.

$$A = \{U, D, L, R\}$$

(b)

$$P_D = \begin{bmatrix}
0,2 & 0 & 0 & 0,8 & 0 & 0 & 0 \\
0 & 0,2 & 0 & 0 & 0 & 0,8 & 0 \\
0 & 0 & 0,2 & 0 & 0 & 0 & 0,8 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

$$C(x,a) = [C]_{xa}$$

$$C = \begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1
\end{bmatrix}$$

(C)

$$\pi = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} ; \quad C_\pi = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} ; \quad P_\pi = P_D$$

$$\gamma = 0{,}9$$

$$J^\pi = (I - \gamma P_\pi)^{-1} \cdot C_\pi$$

Using the numpy library for python, we obtain

$$J^\pi = \begin{bmatrix} 10 \\ 8{,}78 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}$$