

# Planing, Learning and Intelligent Decision Making - Homework 3

99326 - Sebastião Carvalho, 99331 - Tiago Antunes

March 9, 2024

## Contents

<b>1 Question 1</b>	<b>1</b>
1 a) . . . . .	1
1 b) . . . . .	1
1 c) . . . . .	3

## 1 Question 1

### 1 a)

Considering  $\mathcal{X}$  as our state space,  $\mathcal{X} = \{1, 2a, 2b, 3, 4\}$ , corresponding to each of the nodes of the graph.

Our action space is  $\mathcal{A} = \{a, b, c\}$ .

Since the observation at each step is the number in the state designation, our action space will be,  $\mathcal{Z} = \{1, 2, 3, 4\}$ ,

### 1 b)

The transition probabilities matrices, given by the edges in the graph are

$$P_a = \begin{matrix} & \begin{matrix} 1 & 2a & 2b & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2a \\ 2b \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{matrix}$$

$$P_b = \begin{matrix} & \begin{matrix} 1 & 2a & 2b & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2a \\ 2b \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0.5 & 0.5 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{matrix}$$

$$P_c = \begin{matrix} & \begin{matrix} 1 & 2a & 2b & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2a \\ 2b \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0.5 & 0.5 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{matrix}$$

The observation probabilities matrices, corresponding to the probability of an observation given the current state and the previous action are

$$O_a = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2a \\ 2b \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}$$

$$O_b = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2a \\ 2b \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

$$O_c = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2a \\ 2b \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

Note how the probability of an observation that would occur in some states is 0, due to it being impossible to transition to that state, given that action.

Now the only thing missing is the immediate cost function, which we can obtain from the costs indicated in the edges of the graph.

The cost matrix for all actions is given by  $C =$

	$a$	$b$	$c$
$1$	$0$	$0$	$0$
$2a$	$0.5$	$1$	$0$
$2b$	$0.5$	$0$	$1$
$3$	$0$	$0$	$0$
$4$	$0$	$0$	$0$

**1 c)**