

Homework 3:

Grupo 27

97040 Rostislav Andreev

97155 João Aparício

$$(a) \quad X = \{1, 2a, 2b, 3, 4\}$$

$$A = \{a, b, c\}$$

$$Z = \{1, 2, 3, 4\}$$

$$(b) \quad P_a = \begin{bmatrix} 0 & 0,5 & 0,5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad P_b = P_c = \begin{bmatrix} 0 & 0,5 & 0,5 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$O_a = O_b = O_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0,5 & 1 & 0 \\ 0,5 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$(c) \quad b_t = [0 \ 0,5 \ 0,5 \ 0 \ 0]$$

After selecting action a:

$$\hat{Q}_{t+1} = b_t P_a = [0 \ 0,5 \ 0,5 \ 0 \ 0] \begin{bmatrix} 0 & 0,5 & 0,5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$= [0 \ 0 \ 0 \ 0,5 \ 0,5]$$

After selecting action b:

$$\hat{Q}_{t+1} = b_t P_b = [0 \ 0,5 \ 0,5 \ 0 \ 0] \begin{bmatrix} 0 & 0,5 & 0,5 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$= [1 \ 0 \ 0 \ 0 \ 0]$$

After selecting action c:

Given that $P_b = P_c$ we get

$$\hat{Q}_{t+1} = b_t P_c = b_t P_b = [1 \ 0 \ 0 \ 0 \ 0]$$