

Trabalho para casa 3 - Partially observable Markov decision problems

1) O espaço de estados X representa a **localização da princesa (torre A ou torre B)**.

O espaço de ações A é constituído pelas ações de **escolher a torre A ou torre B** e pela **ação de observar ("Scope")**.

O espaço de observações Z é constituído pelas observações de **ver a princesa na torre A ou na torre B** e pela observação de **ver nada ("null observation")**.

$$\mathcal{X} = \{A, B\}$$

$$\mathcal{A} = \{A, B, S\}$$

$$\mathcal{Z} = \{A, B, \emptyset\}$$

2) Com base nas informações do enunciado e nos espaços de estados/ações/observações definidos (usando a mesma ordem), as transition probabilities P e as observation probabilities O são dadas pelas seguintes matrizes.

$$P_A = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

$$P_B = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

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

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

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

$$O_S = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.9 & 0 \end{bmatrix}$$

Dado que o objetivo é minimizar o custo de salvar a princesa, a cost function C é a seguinte:

$$C = \begin{bmatrix} 0 & 1 & 0.5 \\ 1 & 0 & 0.5 \end{bmatrix}$$

3) Supondo que o cavaleiro acredita que a princesa está na torre A com probabilidade = 0.7, que decide observar (**ação "Scope"**) e que observou a princesa na torre B, então aplicando a belief update rule, o belief resultante é dado pela seguinte expressão:

$$b_{new} = \xi b_{old} P_S \text{diag}(O_S)$$

$$b_{new} = \xi \begin{bmatrix} 0.7 & 0.3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.1 & 0 \\ 0 & 0.9 \end{bmatrix}$$

$$b_{new} = \begin{bmatrix} 0.21 & 0.79 \end{bmatrix}$$

O símbolo ξ representa a constante de normalização.