ADI | Homework 3.
Partially observable Markov decision problems

Exercise 1

a)

χ = [A, B]

- A = Princess in Tower A;
- B = Princess in Tower B;

A = [a, b, p]

- a = invade Tower A;
- b = invade Tower B;
- p = peer;

Z = χ = [A, B]

- A = Observe Princess in Tower A;
- B = Observe Princess in Tower B;
- N = Nothing is observed;

b)

$P_a = P_b = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ $\quad$ $P_p = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$O_a = O_b = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ $\quad$ $O_p = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.9 & 0 \end{bmatrix}$

$C = \begin{bmatrix} 0 & 1 & 0.2 \\ 1 & 0 & 0.2 \end{bmatrix}$

c) Since there is no initial observation, we have:

$$\alpha^T_0 = [0.7 \quad 0.3]$$

A standard forward computation yields:

$$\alpha^T_1 = \alpha^T_0 \,.\, P_p \,.\, \mathrm{diag}(O_{P\,:,B}) = [0.7 \quad 0.3] \,.\, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} . \begin{bmatrix} 0.1 & 0 \\ 0 & 0.9 \end{bmatrix} = [0.07 \quad 0.27]$$

Finally, upon normalizing, we get:

$$\mu_{1|0:1} = [0.206 \quad 0.794]$$