# Homework 3 – POMDP

Group 27

Pedro Guerreiro - 78264

Pedro Duarte - 78328

---

a) Identify the state space, X, the action space A, and the observation space, Z. You should explicitly model the fact that, when the agent does not peek, it sees nothing.

- **X = {C, D}**, corresponding to the opponent holding the *Ace of Clubs* and the *Ace of Diamonds*, respectively.
- **A = {P, Gc, Gd}**, corresponding to the agent *peeking*, *guessing Ace of Clubs*, and *guessing Ace of Diamonds*, respectively.
- **Z = {Oc, Od, N}**, corresponding to the agent observing an *Ace of Clubs* and an *Ace of Diamonds*, and Nothing respectively.

---

b) Write down the transition probabilities, the observation probabilities and the cost function for this problem. Make sure that the values in your cost function all lie in the interval [0; 1], while respecting the value-relation between actions induced by the rules of the game.

$$P_P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad P_{Gc} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \quad P_{Gd} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

$$O_P = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.9 & 0 \end{bmatrix} \qquad O_{Gc} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad O_{Gd} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

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

---

c) Suppose that, at some time step t, the agent believes that the opponent has the ace of clubs with a probability 0.7, decides to peek and observes an ace of diamonds. Compute the resulting belief.

$$b_t = \begin{bmatrix} 0.7 & 0.3 \end{bmatrix}$$

$$\boldsymbol{b}_{t+1} = \frac{\boldsymbol{b}_t \mathsf{P}_a \mathrm{diag}(\mathsf{O}_{a,z})}{\|\boldsymbol{b}_t \mathsf{P}_a \mathrm{diag}(\mathsf{O}_{a,z})\|_1}$$

$$b_t P_P diag(O_{P,Od}) = \begin{bmatrix} 0.7 & 0.3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.1 & 0 \\ 0 & 0.9 \end{bmatrix}$$

$$b_t P_P diag(O_{P,Od}) = \begin{bmatrix} 0.07 & 0.27 \end{bmatrix}$$

Therefore:

$$b_{t+1} = \begin{bmatrix} 0.21 & 0.79 \end{bmatrix}$$