HOMEWORK 3 – Partially Observable MDP

Exercise 1

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.

- State Space: [‘AC’, ‘AD’]

  AC: Ace of Clubs

  AD: Ace of Diamonds

- Action Space: [‘guessAC’, ‘guessAD’, ‘peek’]
- Observation Space: [‘sawAC’, ‘sawAD’]

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.

- Transition Probabilities:
  - For action guessAC:

```
[[ 0.5  0.5]
 [ 0.5  0.5]]
```

  - For action guessAD:

```
[[ 0.5  0.5]
 [ 0.5  0.5]]
```

  - For action peek:

```
[[1 0]
 [0 1]]
```

- Observation Probabilities:
  - For action guessAC:

```
[[ 0.5  0.5]
 [ 0.5  0.5]]
```

  - For action guessAD:

```
[[ 0.5  0.5]
 [ 0.5  0.5]]
```

Luís Morais 78416
João Rodrigues 78672

- For action peek:

```
[[ 0.9  0.1]
 [ 0.1  0.9]]
```

- Cost Function:

```
[[ 0.   1.   0.1]
 [ 1.   0.   0.1]]
```

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

```
#Exercise 3
alphaZero = np.array([0.7, 0.3])

alphaOne = np.dot(alphaZero, np.diag(pObsPeek[:,1]) )

sumAlpha = np.sum(alphaOne)

alphaOne = alphaOne / sumAlpha
```

Resulting normalized belief:

```
[ 0.20588235  0.79411765]
```

Luís Morais 78416
João Rodrigues 78672