

MSc in Computer Science and Engineering

**Learning and Decision Making 2016-2017**

Homework 3. Partially Observable Markov Decision Problems

**(a)**

X = {A♣, A♦}, the first state means the opponent is holding Ace of Clubs and the second one means he’s holding Ace of Diamonds.

A = {GA♣, GA♦, peek}, the first and second actions are Guess Ace of Clubs and Ace of Diamonds respectively and the third one is peek.

Z = {OA♣, OA♦}, the possible observations are Observe Ace of Clubs and Observe Ace of Diamonds respectively.

**(b)**

O = { = = , = }

P = { = = , = }

C =

**(c)**

The initial belief is that the opponent has an ace of clubs with probability 0.7 which means that the distribution for that belief is [0.7, 0.3].

The agent decides to peek and observes an ace of diamonds, which means that the agent performs the action “peek” and receives the observation OA♦.

Using the forward algorithm:

.

{ OA♦}

=

= diag() \* \* =

Normalizing:

Which means the the belief of the agent after action peek and observation OA♦ is the distribution:

Belief = [0.206, 0.794]