

# Intelligent Multi Agent Systems



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

## Summer Semester 2016, Homework 3 (26 points)

G. Neumann, G. Gebhardt, R. Lioutikov

Due date: Wednesday 20<sup>th</sup> July, 2016, 23:59 CEST (that's 10:59 AM on the following day in Pago Pago!)

### Problem 3.1 Theoretical Questions [0 Points]

- a) POMDP [4 Points]  
1) What is a POMDP?

POMDP stands for Partial Observable Markov Decision Process. In such a Process the real state can not be obtained by the agent but an observation and its confidence. [c.f. 08-PoMDPs p. 4]

- 2) How is it defined? (formal)

A POMDP is defined by:

1. its state space  $s \in S$
2. its action space  $a \in A$
3. its observation space  $o \in O$
4. its transition dynamics  $\Pr(s'|s, a) = P_{s's}^a$
5. its observation probabilities  $O(o|s) = O_{os}$
6. its reward function  $r(s, a) = R_s^a$
7. its initial state probabilities  $\mu_0(s)$
8. the history  $h_t = (o_1, a_1, \dots, a_{t-1}, o_t)$ , that contains all past states and actions

[from 08-PoMDPs p. 5]

- 3) How is it different from a common MDP?

The **true state** is hidden. Thus the agents have to work with observations. For that, all the history of the actions and observations of all the agents have to be saved. (To make assumptions on the real state and thus the expected reward.)

- 4) Given an example of a POMDP

A typical example for this problem is the tiger-door problem. The agent can choose to open one of two doors, behind one of them is a reward and one the other contains punishment. However the agent can only observe the true state with a probability distribution, through listening. [c.f. 08-PoMDPs p. 7]

- b) State Definitions [3 Points]  
1) Name the different state definitions, introduced in the lecture.

1. Information state

## Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

---

### 2. Belief state

2) How do they differ from one another?

*The information state contains the whole history of actions and observations. The belief state describes a probability distribution over states, given the past observations.*

3) What are their advantages/disadvantages?

*The information state allows for explicit evaluation for each time step. But it grows exponentially in the amount of steps ( $T$ ). The belief state is a possibly much shorter description of the true state. For discrete states: In order to compute the Bellman equation we need to discretize the belief state. [c.f. from 08-PoMDPs pp. 10 ff.]*

### c) Alpha Vectors

[3 Points]

1) What are alpha vectors? 2) How do they relate to POMDPs?

*One alpha vector represents the reward from the value function over the state space for one action. The set of alpha vectors that lie on the upper convex hull represent the best possible actions under the current belief state. [c.f. from 08-PoMDPs pp. 24-26]*

3) Which property allows us to use them?

*The reward function of the Bellman equation for the value iteration is linear in the belief state  $b$  ( $b : p(b, a) = \sum_s b(s) r(s, a)$ ). The max-pooling over all actions implies a piecewise linearity in  $b$  for  $V^*(b)$  (convexity).*

### d) Belief-State Value Iteration

[4 Points]

1) Describe Belief-State Value Iteration

*We iterate over the most probable action-observations transitions to obtain the maximal reward in the previous step. This results in the state, we believe to have obtained for the chosen action.*

2) How are alpha vectors used?

*An alpha vector is used to compute the reward function of a belief state for a policy tree. [c.f. 08-PoMDPs p.28]*

3) What are the main update steps?

*There are three update steps:*

1. Get the  $\max_a(s)$  for the action space and the possible thereafter observations  $o'$
2. Averaging over the observations  $o'$  and adding reward leads to the optimal Q values  $Q_{t-1}^*$
3. Consider the best actions ( $\max_a b$ ) to obtain  $V_{t-1}^*$

*[c.f. 08-PoMDPs pp. 29-30]*

---

## Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

---

e) **Point-Based Value Iteration** [4 Points]

1) Describe Point-Based Value Iteration

*Same as the Belief-State Value Iteration, but only rely on a small set of belief states. Updates the belief state in polynomial time.*

2) How are alpha vectors used?

*Each alpha vectors represent a reward "function" for the state space.*

3) What are the main update steps?

1. "Compute alpha vectors for actions and observations"
2. For a subset of belief states, "average over observations, add reward".
3. "Find best action for each belief point."

[c.f. 08-PoMDPs p. 34]

f) **Bayes Adaptive MDPs** [4 Points]

1) What is the main concept of BAMDPs? When are they useful?

*"A Bayes Adaptive MDP defines a POMDP with the MDP model as unobserved state". We put a posterior in the state space. The agent is aware and can define any uncertainty in its partial observed model. [c.f. 08-PoMDPs p. 76]*

2) Why is a Dirichlet distribution a good prior for discrete systems?

*"[T]he Dirichlet distribution is the conjugate prior of the categorical distribution and multinomial distribution." [c.f. [https://en.wikipedia.org/wiki/Dirichlet\\_distribution](https://en.wikipedia.org/wiki/Dirichlet_distribution)] Which fits to our case if the set of states S and set of actions A are discrete (nominal).*

3) When does a BAMDP become a BAPOMDP?

*BAMDPs become BAPOMDPs, when we can measure the true states s and s', but only an observation model  $O_{os} = p(o|s)$ . [c.f. 08-PoMDPs p. 79]*

4) Can we use Belief-State Value Iteration on BAPOMDPs? Is it a good idea?

*The belief state easily becomes huge, as we have to consider the transition history. This makes the Belief-State Value Iteration on BAPOMDPs feasible, only for small problems. [c.f. 08-PoMDPs pp. 79-80]*

## Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

---

g) **Decentralized POMDPs** [4 Points]

- 1) How do they differ from common to POMDPs??

POMDPs consider problems, or more specific the view on the problem in the perspective of a single agent. Decentralized POMDPs, consider partial observable from a point of view with multiple agents. [c.f. 09-DecPoMDPs p. 1]

- 2) How is a Dec-POMDP defined? (formal)

A Dec-POMDP is defined by:

1. it's state space  $s \in S$
2. an action space  $A_i$  for each agent  $i$
3. an observation space  $O_i$  for each agent  $i$
4. the state-action transition dynamics  $P(s'|s, a) = P_{s's}^a$
5. an observation model per agent  $i$ :  $O_i(o|s) = O_{os}^i$
6. a shared reward function for all agents  $I$ :  $r(s, a) = R_s^a$
7. the initial state probabilities  $\mu_0(s)$ .

[c.f. 09-DecPoMDPs p. 4]

- 3) Can we use the same algorithms as for common POMDPs? Why?

We can not use the PoMDP algorithm directly. The state representation of each agent now not only consists of its own observation history, but every possible other agents observation history as well. A local policy is map from an observation history  $o^i = (o_1^i, \dots, o_t^i)$  to actions. Since each agent can only observe the other actions with some certainty, each agent  $i$  have to save a set of possible local policies  $\Pi_i$  for each other agent  $\notin i$ . The set of policies for each other agent is denoted as joint policy  $\langle \pi_1, \dots, \pi_n \rangle$ . [c.f. 09-DecPoMDPs p. 7 and 9] In order to solve a Dec-POMDP we can now perform an exhaustive search over the piecewise product of combinations of the set of joint policies to find the optimal set of local policies. [c.f. 09-DecPoMDPs pp. 12 ff.]

- 4) Are there other ways to solve Dec-POMDPs? Explain one conceptually

The joint policy may be represented as a set of trees. The root and the nodes (as well as the leafs) are the actions, which are connected via the observations (edges). We can evaluate the values of each tree. [c.f. 09-DecPoMDPs pp. 7 and 12] Since there will be many trees that do not contribute to the optimal solutions, we can prune (exclude) these trees. [c.f. 09-DecPoMDPs p. 16] "We can prune a tree  $q_i$  if there is a distribution  $x(q_i^{-1})$  of other trees that has a higher value for all system states and trees of the other agent." [09-DecPoMDPs p. 18]

## Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

4.2 a)

It depends on the case, but most probably Jolly. By the text description we deduce Roberto did not talk at the start, so that means the (B,W,W) state does not occur. Thus it is Filipe's turn and knowing Roberto kept quiet he discards this state and if he sees ~~Jolly~~ with a white foot he can be sure his is black. If he does not speak then Jolly can be sure his foot is black. Thus the game is always solved in maximum 3 steps. By the wording in the text ('all three were quiet') we deduce Jolly saved them with his foot being black.

Given a uniform distribution of the seven possible states the probabilities of each winning are:

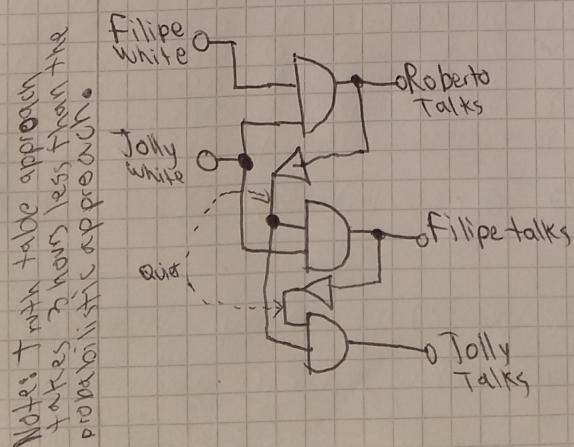
Roberto with Black Foot = 1/7

Filipe with Black Foot = 2/7

Jolly with Black Foot = 4/7

A very simple way for these ruffians to solve this problem is to plug their observations in the WildWest™ circuit to solve their conundrum! As cheap as 49.99 € per piece!

Truth tables:



| F | J | R | T | R' | T' | F' | I | R'' | T'' | F'' |
|---|---|---|---|----|----|----|---|-----|-----|-----|
| 0 | 0 | 0 | 0 | 0  | 0  | 0  | 0 | 0   | 0   | 0   |
| 0 | 1 | 0 | 0 | 0  | 1  | 0  | 0 | 0   | 1   | 0   |
| 1 | 0 | 0 | 0 | 1  | 0  | 0  | 0 | 1   | 0   | 0   |
| 1 | 1 | 1 | 1 | 1  | 1  | 1  | 1 | 1   | 1   | 1   |

| R | F | T | R' | T' | F' |
|---|---|---|----|----|----|
| 0 | 0 | 0 | 0  | 0  | 1  |
| 0 | 0 | 1 | 0  | 1  | 0  |
| 0 | 1 | 0 | 0  | 0  | 0  |
| 0 | 1 | 1 | 1  | 0  | 0  |
| 1 | 0 | 0 | 0  | 0  | 1  |
| 1 | 0 | 1 | 0  | 1  | 0  |
| 1 | 1 | 0 | 1  | 0  | 0  |
| 1 | 1 | 1 | 0  | 0  | 1  |

# Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

Basically it means that if Roberto sees 2 white foot he knows his is black, thus he speaks. By observing Roberto kept quiet Filipe knows both him and Jolly cannot have white, thus if Jolly has white then Filipe knows his is black and speaks. By observing Filipe kept quiet (and therefore also Roberto) Jolly can be certain his color is black, because if his wasn't black either man would have spoken out. The game will never last 4 turns.

Now for the IMAS formulation!

$s_0 \quad s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5 \quad s_6$

$$S = \{B, B, B; B, B, W; B, W, B; B, W, W; W, B, B; W, B, W; W, W, B\}$$

$\Rightarrow 7$  possible states. (Considering the game is over once a solution is found or a false guess leads to an ending.)

- $a \{ \text{stay quiet, talk} \}$  can either guess right or wrong same for all 3 agents

Alternative: •  $a \{ \text{stay quiet, say black, say white} \}$

Not needed

- $O \{ \text{Filipe's foot, Jolly's foot, Jolly Kept Quiet} \}$

Note, If the previous Kept Quiet that assures the one before also did, thus no use within extra observations

$O \{ \text{Jolly's foot, Roberto Kept Quiet} \}$

$O \{ \text{filipe kept quiet} \}$

- $P(s'|s, a) = 1$ . The states don't change what we change is our belief of a state being the true underlying state and we win when we are certain that state is the true state

Roberto only

$$\begin{aligned} O(FF=W|s_0) &= 0 \\ O(FF=W|s_1) &= 0 \\ O(FF=W|s_2) &= 1 \\ O(FF=W|s_3) &= 1 \\ O(FF=W|s_4) &= 0 \\ O(FF=W|s_5) &= 0 \\ O(FF=W|s_6) &= 1 \end{aligned}$$

$$\begin{aligned} O(JF=W|s_0) &= 0 \\ O(JF=W|s_1) &= 1 \\ O(JF=W|s_2) &= 0 \\ O(JF=W|s_3) &= 1 \\ O(JF=W|s_4) &= 0 \\ O(JF=W|s_5) &= 1 \\ O(JF=W|s_6) &= 0 \end{aligned}$$

$$\begin{aligned} O(RKQ=Yes|s_0) &= 1 \\ O(RKQ=Yes|s_1) &= 1 \\ O(RKQ=Yes|s_2) &= 1 \\ O(RKQ=Yes|s_3) &= 0 \\ O(RKQ=Yes|s_4) &= 1 \\ O(RKQ=Yes|s_5) &= 1 \\ O(RKQ=Yes|s_6) &= 1 \end{aligned}$$

$$\begin{aligned} O(FKQ=Yes|s_0) &= 1/7 \\ O(FKQ=Yes|s_1) &= 0 \\ O(FKQ=Yes|s_2) &= 1/7 \\ O(FKQ=Yes|s_3) &= 1/7 \\ O(FKQ=Yes|s_4) &= 2/7 \\ O(FKQ=Yes|s_5) &= 1/7 \\ O(FKQ=Yes|s_6) &= 1/7 \end{aligned}$$

$$\Rightarrow O(FF=W) \approx 3/7 \quad O(JF=W) = 3/7 \quad O(RKQ=Yes) = 6/7 \quad O(FKQ=Yes) = 5/7$$

Given uniform distribution of states

! Do note while  $S$  is supposed to be  $\mathbb{Z}_4$  in reality Filipe can only have 6 possible states where he gets to have a choice and Jolly only gets 4 states

## Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

It kind of fits into the Dec POMDP definition because there are multiple agents in a partially observable world where every agent has their actions and their observations. Every agent (Tally, Filipe and Roberto) has to decide by themselves without communicating directly (decentralized). It does differ slightly because while they share a common goal (stay alive) they also get to keep the foot if they answer (weighted though). They also indirectly know what the others observed through assumptions based on their actions.

History (given longest case) Ex. So

$$H^t = \underbrace{(FF=B, JF=B)}_{\text{OR } LTKQ \text{ distribution}}; \underbrace{a_P = \text{Keep Quiet}}_{\text{DF}}; \underbrace{JF=B, RKQ=\text{Yes}}_{\text{OT}}; \underbrace{a_F = \text{Keep Quiet}}_{\text{DF}}; \underbrace{RKQ=\text{Yes}}_{\text{OT}}; \underbrace{JF=B}_{\text{OT}}$$

b)  $O_i^{x_i V}$  represents the visual (hence v) observations. It contains for Roberto the colour of Filipe and Tally's foot and for Filipe the colour of Tally's foot. Tally gets no visual cues.

$O_i^{x_i A}$  shockingly represents the Auditive observations (yes it was quite clear). For Filipe it contains whether Roberto kept quiet or not (if he talked the game is over though) and for Tally whether Filipe kept quiet (Roberto's is implied anyways)

The visual changes depending on the agent ( $x$ ) and the auditive depending on the time ( $i$ )

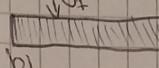
It shows the probabilities of an agent  $x$  observing  $O$  observations based on the states the agent is in. Basically shows that each agent will have corresponding observations and given these probabilities an agent can infer the state by their observations

Suppose  $O_{xV} = FF$  &  $JF$     $O_{xA} = JKQ$  then we have  $\begin{matrix} O_{xV} \\ O_{xA} \end{matrix} = \begin{matrix} FF \\ JKQ \end{matrix}$

$$\begin{matrix} O_{xV} \\ O_{xA} \end{matrix} = \begin{matrix} FF \\ JKQ \end{matrix}$$

# Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

- because the belief state between strokes is correlated, that is there is a true underlying state (just one) and we believe with different certainty which is the most probable. If we start to believe in one more we believe in another less. ~~and~~ this allows nice representations ex . The belief thus depends on the observations

$$\begin{array}{llll} \text{b}_1^R & \text{b}_2^R & \text{b}_3^R & \text{b}_4^R \\ \text{(BB | BB)} = 0.5 & \text{(BB | BW)} = 0 & \text{(BW | BW)} = 0 & \text{(BW | WW)} = 0 \\ \text{(S1 | BB)} = 0 & \text{(S1 | BW)} = 0.5 & \text{(S1 | WW)} = 0 & \text{(S1 | WW)} = 0 \\ \text{(S2 | BB)} = 0 & \text{(S2 | BW)} = 0 & \text{(S2 | BB)} = 0.5 & \text{(S2 | WW)} = 0 \\ \text{(S3 | BB)} = 0 & \text{(S3 | BW)} = 0 & \text{(S3 | BB)} = 0 & \text{(S3 | WW)} = 1 \\ \text{(S4 | BB)} = 0.5 & \text{(S4 | BW)} = 0 & \text{(S4 | BB)} = 0 & \text{(S4 | WW)} = 0 \\ \text{(S5 | BB)} = 0 & \text{(S5 | BW)} = 0.5 & \text{(S5 | BB)} = 0 & \text{(S5 | WW)} = 0 \\ \text{(S6 | BB)} = 0 & \text{(S6 | BW)} = 0 & \text{(S6 | BW)} = 0.5 & \text{(S6 | WW)} = 0 \end{array}$$

|    | BB  | BW  | WW  |
|----|-----|-----|-----|
| S0 | 0.5 | 0   | 0   |
| S1 | 0   | 0.5 | 0   |
| S2 | 0   | 0   | 0.5 |
| S3 | 0   | 0   | 1   |
| S4 | 0.5 | 0   | 0   |
| S5 | 0   | 0.5 | 0   |
| S6 | 0   | 0   | 0.5 |

- It is a Greedy policy given that the reward for choosing wrong (death) is way bigger than the reward for picking the lucky foot to ensure it is only taken when the certainty is 100%. Ex  $r(\text{wrong choice}) = -100000$ ,  $r(\text{right choice}) = 1$ ,  $r(\text{quiet}) = 0$

$$\begin{array}{ll} \text{TR(BB)} = \text{Keep Quiet} & \text{DPA} \\ \text{TR(BW)} = \text{Keep Quiet} & \\ \text{TR(NB)} = \text{Keep Quiet} & \\ \text{TR(WW)} = \text{Say black} & \end{array}$$

0.5 is not enough certainty! we want 100% and 0.1 is the only state with that certainty

- Given it says they all kept quiet he chose to keep quiet (duh) he saw that either (or both) of them had a black foot so he could not be 100% of the colour of his foot and decided to shout up. had he observed two whites he would have said his was black and won
- $\pi$  given the observation tries to estimate  $\pi$  given the state. as we can only guess with uncertainty at the true state  $\pi_0$  is what we use. if we knew our state for certain we'd use  $\pi_S$ .

- If our belief of  $(S1|0) = 1$  then we know with 100% certainty and then  $\pi_{(S1|0)} = \pi_{(S1|0)}$ . If not we use the amount of belief we have

$$\therefore \pi_{(S1|0)} = \pi_{(S1|0)} b_{(S1|0)}$$

# Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

- $\Pi^R(S_0) = \text{keep quiet say black}$
  - $\Pi^R(S_1) = \text{keep quiet say black}$
  - $\Pi^R(S_2) = \text{keep quiet say black}$
  - $\Pi^R(S_3) = \text{say black}$
  - $\Pi^R(S_4) = \text{keep quiet say white}$
  - $\Pi^R(S_5) = \text{keep quiet say white}$
  - $\Pi^R(S_6) = \text{keep quiet say white}$
- If we know the state we can choose with 100% certainty the right action
- |                     |                          |  |
|---------------------|--------------------------|--|
| $\Pi(S_0   KA) = 0$ | $\Pi(S_0   SW) = -10000$ | $\Pi(\text{say black}   S_0) = 1$      |
| $\Pi(S_1   KA) = 0$ | $\Pi(S_1   SW) = -10000$ | $\Pi(\text{say black}   S_1) = 1$      |
| $\Pi(S_2   KA) = 0$ | $\Pi(S_2   SW) = -10000$ | $\Pi(\text{say black}   S_2) = 1$      |
| $\Pi(S_3   KA) = 0$ | $\Pi(S_3   SW) = -10000$ | $\Pi(\text{say black}   S_3) = 1$      |
| $\Pi(S_4   KA) = 0$ | $\Pi(S_4   SW) = 1$      | $\Pi(\text{say black}   S_4) = -10000$ |
| $\Pi(S_5   KA) = 0$ | $\Pi(S_5   SW) = 1$      | $\Pi(\text{say black}   S_5) = -10000$ |
| $\Pi(S_6   KA) = 0$ | $\Pi(S_6   SW) = 1$      | $\Pi(\text{say black}   S_6) = -10000$ |

1 = Yes

0 = No

| JFRKQ            |      | If Roberto talked the game is over and they won or are dead |    |                  |     |                   |    |
|------------------|------|---|----|------------------|-----|-------------------|----|
| $b_2^F(S_0   B)$ | 0.25 | $b_2^F(S_0   B)$  | NA | $b_2^F(S_0   W)$ | 0   | $b_2^F(S_0   WA)$ | NA |
| $(S_1   B)$      | 0    | $(S_1   B)$   | NA | $(S_1   W)$      | 0.5 | $(S_1   WA)$      | NA |
| $(S_2   B)$      | 0.25 | $(S_2   B)$   | NA | $(S_2   W)$      | 0   | $(S_2   WA)$      | NA |
| $(S_3   B)$      | NA   | $(S_3   B)$   | NA | $(S_3   W)$      | NA  | $(S_3   WA)$      | NA |
| $(S_4   B)$      | 0.25 | $(S_4   B)$   | NA | $(S_4   W)$      | 0   | $(S_4   WA)$      | NA |
| $(S_5   B)$      | 0    | $(S_5   B)$   | NA | $(S_5   W)$      | 0.5 | $(S_5   WA)$      | NA |
| $(S_6   B)$      | 0.25 | $(S_6   B)$   | NA | $(S_6   W)$      | 0   | $(S_6   WA)$      | NA |

From this point I will ignore impossible cases (Ex. RT) S1 and S5 both have black foot

| $b_2^F(B)$ | $W$  |
|------------|------|
| $S_0$      | 0.25 |
| $S_1$      | 0    |
| $S_2$      | 0.25 |
| $S_4$      | 0.25 |
| $S_5$      | 0    |
| $S_6$      | 0.25 |

- $\Pi(B)$  Keep Quiet

$\Pi(W)$  Say black

- If we again read it then we see he indeed observed that Roberto kept quiet and one of Tilly's foot is black thus he is 66% certain he has a black which is not safe enough and decides to again shut up

- It informed Filipe that State  $S_3$  ( $S_0, B, W, W$ ) did not occur, because if it was the case the Roberto would have talked. By eliminating this state we can make it a don't care condition which simplifies our state space and modifies directly our beliefs towards more certainty, freeing some information that allows Filipe to 100% know he has the black colour if Filipe has white because thanks to the RKA he knows both can not have white.

- $\Pi^F(S_0) = \text{say black}$
- $\Pi^F(S_1) = \text{say black}$
- $\Pi^F(S_2) = \text{say white}$
- $\Pi^F(S_4) = \text{say black}$
- $\Pi^F(S_5) = \text{say black}$
- $\Pi^F(S_6) = \text{say white}$

|                                    |                                    |
|------------------------------------|------------------------------------|
| $\Pi(S_0   KA) = 0.5$              | $\rightarrow$ all states           |
| $\Pi(\text{Black}   S_0) = 1$      | $\Pi(\text{white}   S_0) = 10000$  |
| $\Pi(\text{Black}   S_1) = 1$      | $\Pi(\text{white}   S_1) = 10000$  |
| $\Pi(\text{Black}   S_2) = -10000$ | $\Pi(\text{white}   S_2) = 1$      |
| $\Pi(\text{Black}   S_4) = 1$      | $\Pi(\text{white}   S_4) = -10000$ |
| $\Pi(\text{Black}   S_5) = 1$      | $\Pi(\text{white}   S_5) = 10000$  |
| $\Pi(\text{Black}   S_6) = -10000$ | $\Pi(\text{white}   S_6) = 1$      |

## Intelligent Multi Agent Systems - Homework 3

Name, Vorname: \_\_\_\_\_ Matrikelnummer: \_\_\_\_\_

Note, Tolly cannot have  $S_4$  nor  $S_5$  nor  $S_6$  nor FT (All tables with OS/KQ)

- $B_3(S_0|FKQ) = 0.25$
- $(S_2|FKQ) = 0.25$
- $(S_4|FKQ) = 0.25$
- $(S_6|FKQ) = 0.25$

|       | FKQ  |
|-------|------|
| $S_0$ | 0.25 |
| $S_2$ | 0.25 |
| $S_4$ | 0.25 |
| $S_6$ | 0.25 |

All these states have Tolly's feet as black

- $\pi^T = \text{Say Black}$ . Tolly should always say black. The game cannot reach Tolly's turn without it being black.

In this formulation

- Tolly thus chose to say his foot was black, saved the three musketeers (chwart!) and kept the foot. He did this by observing that if Roberto didn't talk (which means they both didn't have White) and ~~and~~ Filipe also didn't talk (which in turn means that Tolly didn't have a white, remember, if Tolly would have had white then Filipe could be sure his was black and would talk) by this Tolly can know with certainty that his would be black. By setting the whites to only 2 foot Geraldo shot himself in the foot and ensured the musketeers could always win given they were bright (or took some IMAS class). The game is set up so that by state space reduction once it gets to Tolly it is essentially solved due to elimination of conflicting states

$$\begin{aligned} \pi^T(S_0) &= 1 & \pi^T(\text{Black}|S_0) &= 1 & \pi^T(\text{White}|S_0) &= 0 \\ \pi^T(S_2) &= \{ \text{say black} & \pi^T(\text{Black}|S_2) &= 1 & \pi^T(\text{White}|S_2) &= 0 \\ \pi^T(S_4) &= & \pi^T(\text{Black}|S_4) &= 1 & \pi^T(\text{White}|S_4) &= 0 \\ \pi^T(S_6) &= & \pi^T(\text{Black}|S_6) &= 1 & \pi^T(\text{White}|S_6) &= 0 \end{aligned}$$

- There is never a 4th step, no matter the initial state. If we begin with  $S_3$  then at  $t=1$  Roberto solves it successfully. If we begin at  $S_1$  or  $S_5$  Filipe solves it successfully at  $t=2$ , if we have all the others then Tolly solves it at  $t=3$ . The reason why is in these 6 pages, to make it go for more we would need 3 black 3 white feet