Homework 01

Mariana Ávalos Arce
0197495

Monday, March 7, 2022 7:28 AM

(1) Build the graph of the world defined by the following transition function $f_{MT}(s,a)$:

$$f_{MT}(s,a) = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array} \begin{array}{c} \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} s_2 & s_2 \\ s_1 & s_3 \\ s_3 & s_1 \\ s_1 & s_4 \end{bmatrix} \end{array}$$

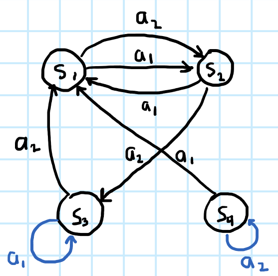
Solution

$$f_{MT}(s,a) = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array} \begin{array}{c} \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} s_2 & s_2 \\ s_1 & s_3 \\ s_3 & s_1 \\ s_1 & s_4 \end{bmatrix} \end{array} \rightarrow \text{deterministic transition function} \longrightarrow s_f = f_{MT}(s,a)$$

Thus,

$$S = \{s_1, s_2, s_3, s_4\} \qquad A = \{a_1, a_2\}$$

The world:



$$\text{with } f_{MT}(s,a) = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array} \begin{array}{c} \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} s_2 & s_2 \\ s_1 & s_3 \\ s_3 & s_1 \\ s_1 & s_4 \end{bmatrix} \end{array}$$

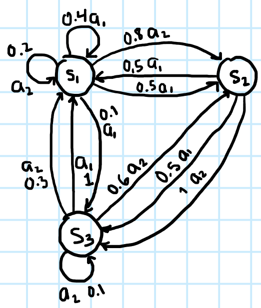
(2) Build the graph of the world defined by the following transition function $P_{MT}(s_f \mid s,a)$:

$$f_{MT}(s,a) = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} s_f = s_1 \\ \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} 0.4 & 0.2 \\ 0.5 & 0 \\ 1 & 0.3 \end{bmatrix} \end{array} \quad \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} s_f = s_2 \\ \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} 0.5 & 0.8 \\ 0 & 0 \\ 0 & 0.6 \end{bmatrix} \end{array} \quad \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} s_f = s_3 \\ \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} 0.1 & 0 \\ 0.5 & 1 \\ 0 & 0.1 \end{bmatrix} \end{array}$$

Solution

Non-deterministic transition function $\quad s_f \sim P_{MT}(s_f \mid s, a)$

Thus, the world looks like below:



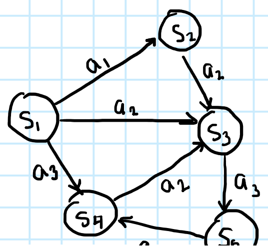
sum = 1

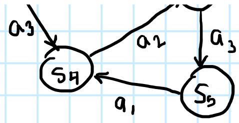
$$\text{with } f_{MT}(s,a) = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} s_f = s_1 \\ \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} 0.4 & 0.2 \\ 0.5 & 0 \\ 1 & 0.3 \end{bmatrix} \end{array} \quad \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} s_f = s_2 \\ \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} 0.5 & 0.8 \\ 0 & 0 \\ 0 & 0.6 \end{bmatrix} \end{array} \quad \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} s_f = s_3 \\ \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} 0.1 & 0 \\ 0.5 & 1 \\ 0 & 0.1 \end{bmatrix} \end{array}$$

this means no self transition

no a_1 action makes S_3 stay in itself

(3) Build the transition function $f_{MT}(s,a)$ of the following world:





Solution

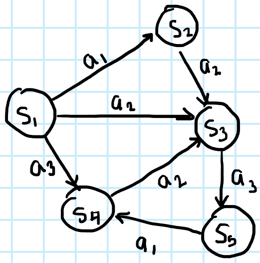
We can define $S=\{s_1, s_2, s_3, s_4, s_5\}$ and $A=\{a_1, a_2, a_3\}$, and by looking at the world, the transition function is deterministic:

$$s_f = f_{MT}(s,a)$$

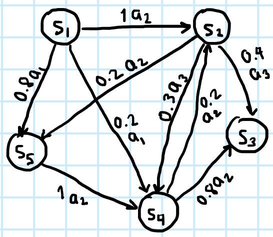
Therefore,

$$s_f = f_{MT}(s,a) = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} s_2 & s_3 & s_4 \\ s_2 & s_3 & s_2 \\ s_3 & s_3 & s_5 \\ s_4 & s_3 & s_4 \\ s_4 & s_5 & s_5 \end{bmatrix} \end{array}$$

for:



④ Build the transition function $P_{MT}(s_f \mid s,a)$ for the world:

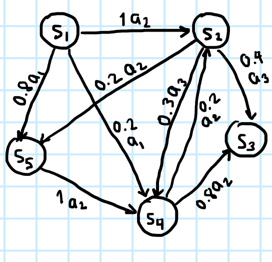


Solution:

By looking at the world, it needs a non-deterministic transition function:

$S=\{s_1, s_2, s_3, s_4, s_5\}$ and $A=\{a_1, a_2, a_3\}$

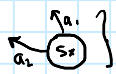
$$s_f \sim P_{MT}(s_f \mid s,a) = \overset{s_f = s_1}{\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0.2 & 0.1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}} \overset{s_f = s_2}{\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0.2 & 0.1 \\ 0 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}} \overset{s_f = s_3}{\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.2 & 0.4 \\ 1 & 1 & 1 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}} \overset{s_f = s_4}{\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0.3 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \end{array}} \overset{s_f = s_5}{\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 0.2 & 0.1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array}}$$



Notes to myself:

Steps:

1. Fill all blacks by going matrix by matrix: Which ones arrive to this matrix?
2. fill every $s_i a_j$ remaining with its probability left, two cases:
   ↓ if they got blacks
   a) No prob left: zeros
   b) Prob left: i.e. blacks are $0.4 + 0.3 \rightarrow 0.3$ left to 3 matrices, 0.1 for each
3. If all matrices' row x is empty: go to S_x row for matrix $S_f = S_x$ and:



$\}$ no arriving arrow = fill S_x row for $S_f = S_x$ with 1's.

4. For all remaining $s_i a_j$ cells in all its matrices, look for $\underbrace{S_f = S_i \text{ and row } S_i}_{\text{row you're checking}}$, put 1. For ex, S_5 for $S_f = S_5$

⑤ The world has the following set of states $S=\{s_1, s_2, s_3, SF1, SF2\}$ where $s1$ = initial state and, SF1 an SF2 are terminal states:

sF1	s1	s2	s3	sF2

The world has the following set of actions $A=\{\rightarrow, \leftarrow\}$

The reward function $f_R(s,a,s_f) = f_R(s_f)$ only depends on the state to which the agent arrives and is defined as:

-10	0	-0.4	-0.4	10

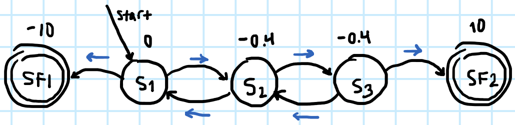
That is, if the agent goes from s_1 to s_2 then it receives -0.4 of reward that is defined in state agent arrives and is defined as:

-10	0	-0.4	-0.4	10

That is, if the agent goes from s_1 to s_2 then it receives -0.4 of reward that is defined in state s_2.

Do as follows:

a) Build the graph of the world:



b) Write the transition function $f_{MT}(s,a)$

$$f_{MT}(s,a) = \begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & SF_1 & SF_1 \\ s_1 & s_2 & SF_1 \\ s_2 & s_3 & s_1 \\ s_3 & SF_2 & s_2 \\ SF_2 & SF_2 & SF_2 \end{array}$$

c) Write the reward function $f_R(s,a,s_f) = f_R(s_f)$

Solution

$$f_R(s,a,s_f) = \underbrace{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & -10 & -10 \\ s_1 & -10 & -10 \\ s_2 & -10 & -10 \\ s_3 & -10 & -10 \\ SF_2 & -10 & -10 \end{array}}_{s_f = SF_1} \quad \underbrace{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 0 & 0 \\ s_1 & 0 & 0 \\ s_2 & 0 & 0 \\ s_3 & 0 & 0 \\ SF_2 & 0 & 0 \end{array}}_{s_f = s_1} \quad \underbrace{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & -0.4 & -0.4 \\ s_1 & -0.4 & -0.4 \\ s_2 & -0.4 & -0.4 \\ s_3 & -0.4 & -0.4 \\ SF_2 & -0.4 & -0.4 \end{array}}_{s_f = s_2} \quad \underbrace{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & -0.4 & -0.4 \\ s_1 & -0.4 & -0.4 \\ s_2 & 0.4 & -0.4 \\ s_3 & -0.4 & -0.4 \\ SF_2 & -0.4 & -0.4 \end{array}}_{s_f = s_3} \quad \underbrace{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 10 & 10 \\ s_1 & 10 & 10 \\ s_2 & 10 & 10 \\ s_3 & 10 & 10 \\ SF_2 & 10 & 10 \end{array}}_{s_f = SF_2}$$

$$= f_R(s_f) = \underset{s_f = SF_1}{s\overset{a}{\begin{bmatrix} -10 \end{bmatrix}}} \quad \underset{s_f = SF_1}{s\overset{a}{\begin{bmatrix} 0 \end{bmatrix}}} \quad \underset{s_f = SF_1}{s\overset{a}{\begin{bmatrix} -0.4 \end{bmatrix}}} \quad \underset{s_f = SF_1}{s\overset{a}{\begin{bmatrix} -0.4 \end{bmatrix}}} \quad \underset{s_f = SF_1}{s\overset{a}{\begin{bmatrix} 10 \end{bmatrix}}} = s_f \begin{array}{c} SF_1 \\ s_1 \\ s_2 \\ s_3 \\ SF_2 \end{array} \begin{bmatrix} -10 \\ 0 \\ -0.4 \\ -0.4 \\ 10 \end{bmatrix}$$

d) Write two action functions $f_\pi(s)$:

i. One so that the agent, from the initial state, arrives to SF_2 (final state).

Solution

Deterministic action tion $a = f_\pi(s)$. Thus, from the initial state to SF_2,

$$a = f_\pi(s) = \begin{array}{c} SF_1 \\ s_1 \\ s_2 \\ s_3 \\ SF_2 \end{array} \begin{bmatrix} \text{None} \\ \rightarrow \\ \rightarrow \\ \rightarrow \\ \text{None} \end{bmatrix}$$

ii. One so that the agent, from the initial state, arrives to SF_1 (final state).

Solution

Thus, from initial state s_1, to arrive to SF_1 the action function is

$$a = f_\pi(s) = \begin{array}{c} SF_1 \\ s_1 \\ s_2 \\ s_3 \\ SF_2 \end{array} \begin{bmatrix} \text{None} \\ \leftarrow \\ \leftarrow \\ \leftarrow \\ \text{None} \end{bmatrix}$$

6) The world has the following set of states $S = \{s_1, s_2, s_3, SF1, SF2\}$ where $s1$ = initial state and, SF1 an SF2 are terminal states:

sF1	s1	s2	s3	sF2

The world has the following set of actions $A = \{\rightarrow, \leftarrow\}$, where:

S_{F_1}	S_1	S_2	S_3	S_{F_2}

The world has the following set of actions $A = \{\rightarrow, \leftarrow\}$, where:

$\rightarrow$ = The agent moves to the right with probability of 0.8 and to the left with prob. of 0.2

$\leftarrow$ = The agent moves to the left with probability of 0.8 and to the right with prob. of 0.2

The reward function $f_R(s,a,s_f) = f_R(s_f)$ only depends on the state to which the agent arrives and is defined as:

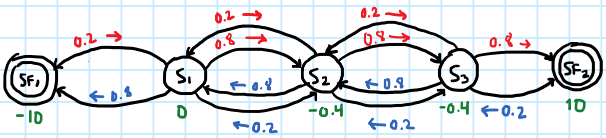
-10	0	-0.4	-0.4	10

That is, if the agent goes from s_1 to s_2 then it receives -0.4 of reward that is defined in state s_2.

Do as follows:

a) Build the graph of the world

Solution

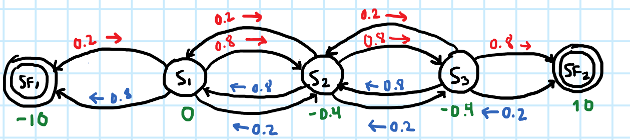


b) Write the transition function $P_{MT}(s_f \mid s, a)$

Solution

The non-deterministic transition function is

$$s_f \sim P_{MT}(s_f \mid s,a) = \underset{s_f = SF_1}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 1 & 1 \\ s_1 & 0.2 & 0.8 \\ s_2 & 0 & 0 \\ s_3 & 0 & 0 \\ SF_2 & 0 & 0 \end{array}} \quad \underset{s_f = s_1}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 0 & 0 \\ s_1 & 0.2 & 0.05 \\ s_2 & 0.2 & 0.8 \\ s_3 & 0 & 0 \\ SF_2 & 0 & 0 \end{array}} \quad \underset{s_f = s_2}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 0 & 0 \\ s_1 & 0.2 & 0.05 \\ s_2 & 0 & 0 \\ s_3 & 0.2 & 0.8 \\ SF_2 & 0 & 0 \end{array}} \quad \underset{s_f = s_3}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 0 & 0 \\ s_1 & 0.2 & 0.05 \\ s_2 & 0.8 & 0.2 \\ s_3 & 0 & 0 \\ SF_2 & 0 & 0 \end{array}} \quad \underset{s_f = SF_2}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 0 & 0 \\ s_1 & 0.2 & 0.05 \\ s_2 & 0 & 0 \\ s_3 & 0.8 & 0.2 \\ SF_2 & 1 & 1 \end{array}}$$



c) Write the reward function $f_R(s,a,s_f) = f_R(s_f)$

Solution

$$f_R(s,a,s_f) = \underset{s_f = SF_1}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & -10 & -10 \\ s_1 & -10 & -10 \\ s_2 & -10 & -10 \\ s_3 & -10 & -10 \\ SF_2 & -10 & -10 \end{array}} \quad \underset{s_f = s_1}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 0 & 0 \\ s_1 & 0 & 0 \\ s_2 & 0 & 0 \\ s_3 & 0 & 0 \\ SF_2 & 0 & 0 \end{array}} \quad \underset{s_f = s_2}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & -0.4 & -0.4 \\ s_1 & -0.4 & -0.4 \\ s_2 & -0.4 & -0.4 \\ s_3 & -0.4 & -0.4 \\ SF_2 & -0.4 & -0.4 \end{array}} \quad \underset{s_f = s_3}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & -0.4 & -0.4 \\ s_1 & -0.4 & -0.4 \\ s_2 & -0.4 & -0.4 \\ s_3 & -0.4 & -0.4 \\ SF_2 & -0.4 & -0.4 \end{array}} \quad \underset{s_f = SF_2}{\begin{array}{c|cc} & \rightarrow & \leftarrow \\ \hline SF_1 & 10 & 10 \\ s_1 & 10 & 10 \\ s_2 & 10 & 10 \\ s_3 & 10 & 10 \\ SF_2 & 10 & 10 \end{array}}$$

$$= f_R(s_f) = \underset{s_f = SF_1}{s\begin{bmatrix} -10 \end{bmatrix}} \quad \underset{s_f = SF_1}{s\begin{bmatrix} 0 \end{bmatrix}} \quad \underset{s_f = SF_1}{s\begin{bmatrix} -0.4 \end{bmatrix}} \quad \underset{s_f = SF_1}{s\begin{bmatrix} -0.4 \end{bmatrix}} \quad \underset{s_f = SF_1}{s\begin{bmatrix} 10 \end{bmatrix}} = \begin{array}{c|c} SF_1 & -10 \\ s_1 & 0 \\ s_2 & -0.4 \\ s_3 & -0.4 \\ SF_2 & 10 \end{array}$$

d) Write two action functions $f_\pi(s)$:

i. One so that the agent, from the initial state, arrives to SF_2 (final state).

Solution

Deterministic action tion $a = f_\pi(s)$. Thus, from the initial state to SF_2,

$$a = f_\pi(s) = \begin{array}{c|c} SF_1 & \text{None} \\ s_1 & \rightarrow \\ s_2 & \rightarrow \\ s_3 & \rightarrow \\ SF_2 & \text{None} \end{array}$$

ii. One so that the agent, from the initial state, arrives to SF_1 (final state).

Solution

Thus, from initial state S_1, to arrive to SF_1 the action function is

$$a = f_{\pi}(s) = SF_1 \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ SF_2 \end{array} \begin{bmatrix} \text{None} \\ \leftarrow \\ \leftarrow \\ \leftarrow \\ \text{None} \end{bmatrix}$$

e) Show the probability of the agent arriving at the final state sF_2 and at the final state sF_1 in both politics of question d.

Solution

i) To sF_2

$S_1 \rightarrow S_2 \quad p = 0.8$ $\qquad P(S_1 \rightarrow S_2 \mid S_1)$

$S_2 \rightarrow S_3 \quad p = 0.8$ $\qquad P(S_2 \rightarrow S_3 \mid S_2)$

$S_3 \rightarrow SF_2 \quad p = 0.8$ $\qquad P(S_3 \rightarrow SF_2 \mid S_3) = (0.8)(0.8)(0.8)$

$$= \boxed{0.512}$$

ii) To sF_1

$S_1 \rightarrow SF_1 = 0.8$

$$P(S_1 \rightarrow SF_1 \mid S_1) = \boxed{0.8}$$