

# Homework 01

Monday, March 7, 2022 7:28 AM

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

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

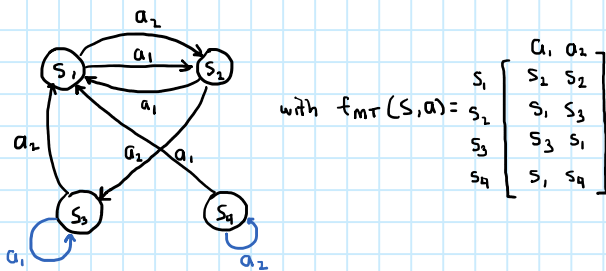
Solution

$$f_{MT}(s, a) = \begin{matrix} & a_1, a_2 \\ s_1 & \begin{bmatrix} s_2 & s_2 \\ s_1 & s_3 \\ s_3 & s_1 \\ s_4 & s_4 \end{bmatrix} \end{matrix} \rightarrow \text{deterministic transition function} \rightarrow s_t = f_{MT}(s, a)$$

Thus,

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

The world:



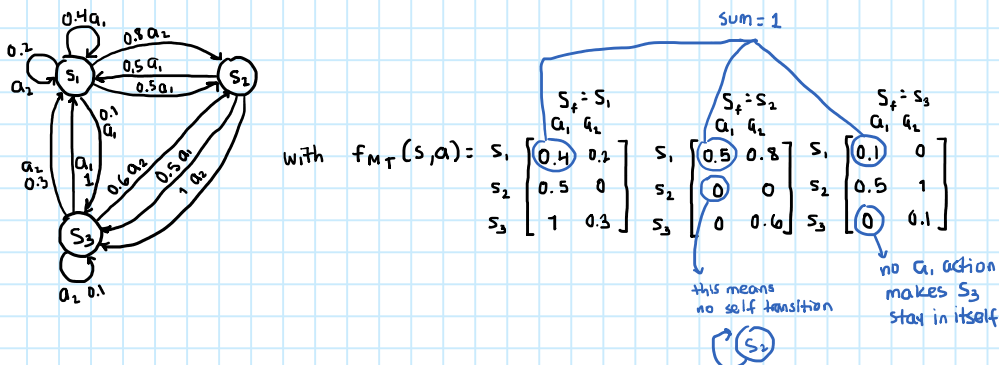
- ② Build the graph of the world defined by the following transition function  $P_{MT}(s_t | s, a)$ :

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

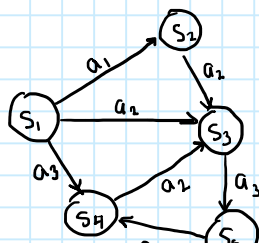
Solution

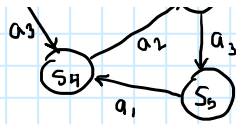
Non-deterministic transition function  $s_t \sim P_{MT}(s_t | s, a)$

Thus, the world looks like below:



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





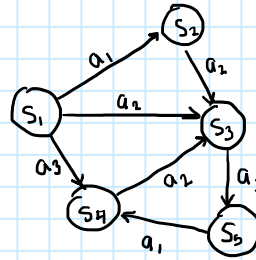
Solution

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:

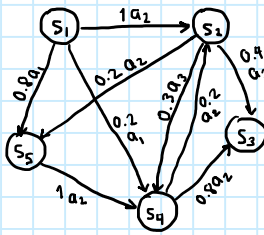
$$S_f = f_{MT}(S, A)$$

Therefore,

$$S_f = f_{MT}(s_i, a) = s_j \quad \begin{matrix} a_1 & a_2 & a_3 \\ s_1 & s_2 & s_3 & s_4 \\ s_2 & s_2 & s_3 & s_2 \\ s_3 & s_3 & s_3 & s_5 \\ s_4 & s_4 & s_3 & s_4 \\ s_5 & s_4 & s_3 & s_5 \end{matrix} \quad \text{for:}$$



④ Build the transition function  $P_{MT}(S_f | s_i, a)$  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  $A = \{a_1, a_2, a_3\}$

$$S_f \sim P_{MT}(S_f | s_i, a) = \begin{matrix} & S_f = s_1 & S_f = s_2 & S_f = s_3 & S_f = s_4 & S_f = s_5 \\ & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} \\ s_1 & \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0.2 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0.8 & 0 & 0 \end{bmatrix} \\ s_2 & \begin{bmatrix} 0 & 0.2 & 0.1 \end{bmatrix} & \begin{bmatrix} 1 & 0.2 & 0.1 \end{bmatrix} & \begin{bmatrix} 0 & 0.2 & 0.4 \end{bmatrix} & \begin{bmatrix} 0 & 0.2 & 0.3 \end{bmatrix} & \begin{bmatrix} 0 & 0.2 & 0.1 \end{bmatrix} \\ s_3 & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \\ s_4 & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0.2 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0.8 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \\ s_5 & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \end{matrix}$$

Notes to myself:

Steps:

1. Fill all blacks by going matrix by matrix: Which ones arrive to this matrix?

2. Fill every  $S_{ij}$  remaining with its probability left, two cases:

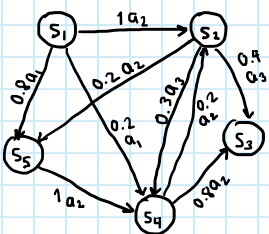
a) No prob left: zeros

b) Prob left: i.e. blacks are  $0.4 + 0.3 \rightarrow 0.3$  left to 3 matrices, 0.1 for each

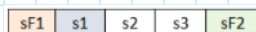
3. If all matrices' row  $x$  is empty: go to  $S_x$  row for matrix  $S_f = S_x$  and:

no arriving arrow = fill  $S_x$  row for  $S_f = S_x$  with 1's.  
row you're checking

4. For all remaining  $S_{ij}$  cells in all its matrices, look for  $S_f = S_i$  and row  $S_i$ , put 1. For ex,  $S_5$  for  $S_f = S_5$

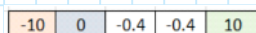


⑤ The world has the following set of states  $S = \{s_1, s_2, s_3, sF1, sF2\}$  where  $s_1$  = initial state and,  $sF1$  or  $sF2$  are terminal states:



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:



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

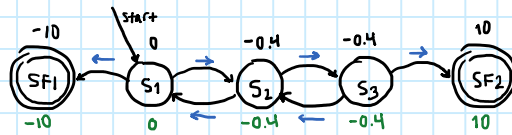
agent arrives and is defined as:

-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  $f_{MT}(s, a)$

$$f_{MT}(s, a) = \begin{matrix} \rightarrow & \leftarrow \\ s_f & s_i \end{matrix} \begin{matrix} SF_1 & SF_1 \\ s_1 & SF_1 \\ s_2 & SF_1 \\ s_3 & SF_1 \\ SF_2 & SF_1 \end{matrix}$$

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

Solution

$$f_R(s, a, s_f) = \begin{matrix} \rightarrow & \leftarrow \\ s_f & s_i \end{matrix} \begin{matrix} SF_1 & SF_1 \\ s_1 & SF_1 \\ s_2 & SF_1 \\ s_3 & SF_1 \\ SF_2 & SF_1 \end{matrix} \begin{matrix} -10 & -10 \\ -10 & -10 \\ -10 & -10 \\ -10 & -10 \\ -10 & -10 \end{matrix}$$

$$= f_R(s_f) = \begin{matrix} a \\ s_f = SF_1 \end{matrix} \begin{bmatrix} -10 \\ 0 \\ -0.4 \\ -0.4 \\ 10 \end{bmatrix} = \begin{matrix} a \\ s_f = SF_1 \end{matrix} \begin{bmatrix} -10 \\ 0 \\ -0.4 \\ -0.4 \\ 10 \end{bmatrix}$$

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

$$a = f_\pi(s) = \begin{matrix} s_f & s_i \end{matrix} \begin{matrix} None \\ \rightarrow \\ \rightarrow \\ \rightarrow \\ None \end{matrix}$$

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

$$a = f_\pi(s) = \begin{matrix} s_f & s_i \end{matrix} \begin{matrix} None \\ \leftarrow \\ \leftarrow \\ \leftarrow \\ None \end{matrix}$$

6) The world has the following set of states  $S = \{s_1, s_2, s_3, SF_1, SF_2\}$  where  $s_1$  = initial state and,  $SF_1$  or  $SF_2$  are terminal states:

SF1	s1	s2	s3	SF2
-----	----	----	----	-----

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

SF1	S1	S2	S3	SF2
-----	----	----	----	-----

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:

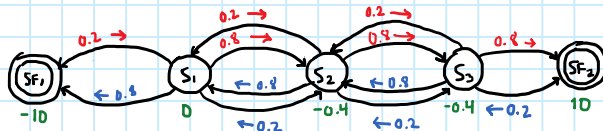
-10	0	-0.4	-0.4	10
-----	---	------	------	----

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

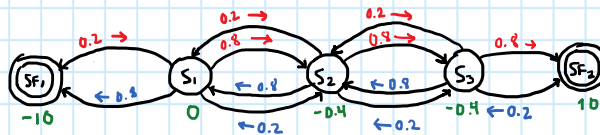


b) Write the transition function  $P_{MT}(s_f | s, a)$

Solution

The non-deterministic transition function is

$$s_f \sim P_{MT}(s_f | s, a) = \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ 1 & 1 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.2 & 0.8 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = SF_1 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.2 & 0.05 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.2 & 0.8 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = S_1 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.1 & 0.05 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.2 & 0.8 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = S_2 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.2 & 0.05 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.8 & 0.2 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = S_3 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.1 & 0.05 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.8 & 0.2 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 1 & 1 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = SF_2 \end{matrix} \end{matrix}$$



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

Solution

$$f_R(s, a, s_f) = \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ -10 & -10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -10 & -10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -10 & -10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -10 & -10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -10 & -10 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = SF_1 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = S_1 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = S_2 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = S_3 \end{matrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ 10 & 10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 10 & 10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 10 & 10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 10 & 10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 10 & 10 \end{matrix} \end{bmatrix} & \begin{matrix} s_f = SF_2 \end{matrix} \end{matrix}$$

$$= f_R(s_f) = \begin{matrix} & \begin{matrix} \rightarrow & \leftarrow \end{matrix} \\ \begin{matrix} s_f \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} & \begin{bmatrix} \begin{matrix} \rightarrow & \leftarrow \\ -10 & -10 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 0 & 0 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ -0.4 & -0.4 \end{matrix} \\ \begin{matrix} \rightarrow & \leftarrow \\ 10 & 10 \end{matrix} \end{bmatrix} = \begin{bmatrix} s_{f1} \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{bmatrix} \begin{bmatrix} -10 \\ 0 \\ -0.4 \\ -0.4 \\ 10 \end{bmatrix}$$

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

$$\pi = f_\pi(s) = \begin{matrix} s_{f1} \\ s_1 \\ s_2 \\ s_3 \\ s_{f2} \end{matrix} \begin{bmatrix} \text{None} \\ \rightarrow \\ \rightarrow \\ \rightarrow \\ \text{None} \end{bmatrix}$$

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

$$a = f_{\pi}(s) = SF_1 \begin{bmatrix} \text{None} \\ \leftarrow \\ \leftarrow \\ \leftarrow \\ \text{None} \end{bmatrix}$$

e) Show the probability of the agent arriving at the final state  $SF_2$  and at the final state  $SF_1$  in both policies of question d.

Solution

i) To  $SF_2$

$$S_1 \rightarrow S_2 \quad p = 0.8$$

$$S_1 \rightarrow S_3 \quad p = 0.8$$

$$S_3 \rightarrow SF_2 \quad p = 0.8$$

$$P(S_1 \rightarrow S_2 | S_1)$$

$$P(S_2 \rightarrow S_3 | S_2)$$

$$P(S_3 \rightarrow SF_2 | S_3) = (0.8)(0.8)(0.8)$$

$$= \boxed{0.512}$$

ii) To  $SF_1$

$$S_1 \rightarrow SF_1 = 0.8$$

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