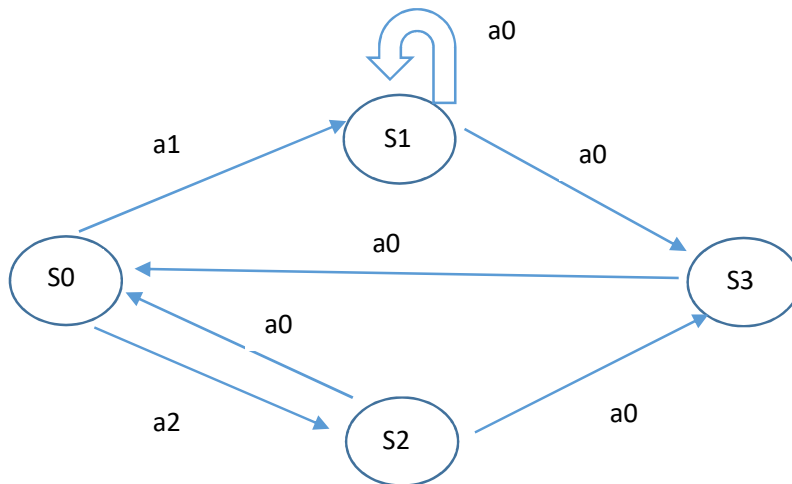


## ROB311 - TP2 – 21 Septembre 2020



In the figure above, the states are depicted by circles (S0, S1, S2, and S3) and the associated actions are indicated on the arrows: a0, a1, and a2. The transition functions for all the actions are shown below.

$$T(S, a0, S') = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1-x & 0 & x \\ 1-y & 0 & 0 & y \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$T(S, a1, S') = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T(S, a2, S') = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Each of the parameters  $x$  and  $y$  are in the interval  $[0, 1]$ , and the discounted factor  $\gamma \in [0, 1)$

The reward is:

$$R(s) = \begin{cases} 10, & \text{for state } S3 \\ 1, & \text{for state } S2 \\ 0, & \text{otherwise} \end{cases}$$

Question 1:

Enumerate all the possible policies

Question 2:

Write the equation for each optimal value function for each state

(  $V^*(s_0), V^*(s_1), V^*(s_2), V^*(s_3)$  )

Reminder:

$$V^*(S) = R(s) + \max_a \gamma \sum_{S'} T(S, a, S') V^*(S')$$

Question 3:

Is there exist a value for x, that for all  $\gamma \in [0,1)$ , and  $y \in [0,1]$  ,  $\pi^*(s_0) = a_2$ . Justify your answer.

Reminder:

$$\pi^*(s) = \arg \max_a \sum_{S'} T(S, a, S') V^*(S')$$

Question 4:

Is there exist a value for y, that for all  $x > 0$ , and  $\gamma \in [0,1]$  ,  $\pi^*(s_0) = a_1$ . Justify your answer.

Question 5:

Using  $x=y=0.25$  and  $\gamma = 0.9$  , calculate the  $\pi^*$  and  $V^*$  for all states.

Implement value iteration.