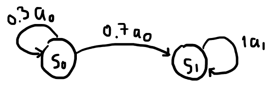
Exam 1: practice

Wednesday, April 6, 2022 9:49 AM

1. 

$s_f \sim P(s_f | s, a) =$

$s_f = s_0$:

	a_0	a_1
s_0	0.3	1
s_1	0	0

$s_f = s_1$:

	a_0	a_1
s_0	0.7	0
s_1	1	1

$f_R(s, a, s_f) =$

$s_f = s_0$:

	a_0	a_1
s_0	0	1
s_1	1	2

$s_f = s_1$:

	a_0	a_1
s_0	-3	-1
s_1	1	0

$a = f_\pi(s) =$

s_0	a_0
s_1	a_1

with $\gamma = 0.7$

Calculate the two $V(S) = \{V(s_0), V(s_1)\}$

$$V(s_0) = \sum_{i=0}^{1} P_{MT}(s_i | s_0, a_0)\left[f_R(s_0, a_0, s_i) + \gamma V(s_i)\right]$$

$$= P_{MT}(s_0 | s_0, a_0)[f_R(s_0, a_0, s_0) + \gamma V(s_0)] +$$
$$P_{MT}(s_1 | s_0, a_0)[f_R(s_0, a_0, s_1) + \gamma V(s_1)]$$

$$V(s_0) = (0.3)[0 + (0.7)V(s_0)] + (0.7)[-3 + (0.7)V(s_1)]$$
$$V(s_0) = 0.21 V(s_0) + [-2.1 + 0.49 V(s_1)]$$
$$1 V(s_0) - 0.21 V(s_0) = -2.1 + 0.49 V(s_1)$$
$$0.79 V(s_0) = -2.1 + 0.49 V(s_1)$$
$$0.79 V(s_0) - 0.49 V(s_1) = -2.1 \qquad (1)$$

$$V(s_1) = \sum_{i=0}^{1} P_{MT}(s_i | s_1, a_1)\left[f_R(s_1, a_1, s_i) + \gamma V(s_i)\right]$$

$$= P_{MT}(s_0 | s_1, a_1)[f_R(s_1, a_1, s_0) + \gamma V(s_0)] +$$
$$P_{MT}(s_1 | s_1, a_1)[f_R(s_1, a_1, s_1) + \gamma V(s_1)]$$

The system is

$0.79 V(s_0) - 0.49 V(s_1) = -2.1 \quad (1)$

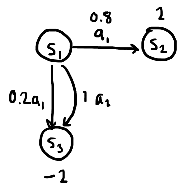
$V(s_1) = 0 \quad (2)$

$$V(s_1) = 0[2 + (0.7)V(s_0)] + 1[0 + 0.7 V(s_1)]$$
$$V(s_1) = 0.7 V(s_1)$$
$$1 V(s_1) - 0.7 V(s_1) = 0$$
$$0.3 V(s_1) = 0$$
$$\boxed{V(s_1) = 0} \qquad (2)$$

Plugging (2) into (1)

$$0.79 V(s_0) - 0.49(0) = -2.1$$
$$0.79 V(s_0) = -2.1$$
$$V(s_0) = \frac{-2.1}{0.79}$$
$$\boxed{V(s_0) = -2.65}$$

2. There's an stochastic world:



$f_R(s_f) =$

s_1	0
s_2	2
s_3	-2

$a = f_\pi(s) =$

s_1	a_1
s_2	a_2
s_3	a_3

} Deterministic, all $p = 1$

From the politic, calculate the average f_{AR} for the trajectories from $f_\pi(s)$

$\tau_1 = s_1 \xrightarrow{0.8 a_1} s_2$ (2) $\qquad \tau_2 = s_1 \xrightarrow{0.2 a_1} s_3$ (-2)

$f_{AR} = 2 \qquad f_{AR} = -2$

$P(\tau_1) = 1(0.8) = 0.8 \qquad P(\tau_2) = 1(0.2) = 0.2$

$$\text{denom} = \frac{\#\ \tau\text{'s I want}}{\text{Total } \tau\text{'s}} = \frac{2}{2} = 1$$

$$\overline{f_{AR}} = \frac{0.8}{1}(2) + \frac{(0.2)}{1}(-2) = 1.6 - 0.4 = \boxed{1.2}$$