ROB311 - TP2

Benjamin Collery and Clément Fauchereau

September 2019

1 Question 1

The possible policies are:

State	S0	S1	S2	S3
Action	a1	a0	a0	a0

and

State	S0	S1	S2	S3
Action	a2	a0	a0	a0

2 Question 2

$$
\begin{aligned}
V^*(S0) &= max(\gamma V^*(S1), \gamma V^*(S2)) \\
(1 + \gamma x - \gamma)V^*(S1) &= x\gamma V^*(S3) \\
V^*(S2) &= 1 + \gamma(1-y)V^*(S0) + \gamma y V^*(S3) \\
V^*(S3) &= 10 + \gamma V^*(S0)
\end{aligned}
$$

3 Question 3

$$\pi^*(S0) = argmax(\gamma V^*(S1), \gamma V^*(S2))$$

With $x = 0$:
$V^*(S1) = \gamma V^*(S1)$ and $\gamma \neq 1$ therefore $V^*(S1) = 0$
$V^*(S2) >= 1$ therefore $V^*(S2) > V^*(S1)$ and $\pi^*(S0) = a2$

4 Question 5

Using $x = y = 0.25$ and $\gamma = 0.9$, the optimal policy is:

State	S0	S1	S2	S3
Action	a1	a0	a0	a0

and utilities are:

State	S0	S1	S2	S3
V^*	14.176	15.752	15.688	22.758