TP2

kai ZHANG, mengyu PAN

September 2020

1 Q1

$$\pi(S_0) : S_0 \to a_1, a_2$$

$$\pi(S_1) : S_1 \to a_0$$

$$\pi(S_2) : S_2 \to a_0$$

$$\pi(S_3) : S_3 \to a_0$$

2 Q2

$$\begin{aligned} V^*(S_0) &= R(S_0) + max_a \gamma \begin{bmatrix} T(S_0, a_1, S_1)V^*(S_1) \\ T(S_0, a_2, S_2)V^*(S_2) \end{bmatrix} \\ &= max_a \gamma \begin{bmatrix} V^*(S_1) \\ V^*(S_2) \end{bmatrix} \end{aligned}$$

$$\begin{aligned} V^*(S_1) &= R(S_1) + max_a \gamma [T(S_1, a_0, S_1)V^*(S_1) + T(S_1, a_0, S_3)V^*(S_3)] \\ &= max_a \gamma [(1-x)V^*(S_1) + xV^*(S_3)] \\ &= \gamma [(1-x)V^*(S_1) + xV^*(S_3)] \\ &= \frac{x\gamma}{1-\gamma+} V^*(S_3) \\ &= \frac{x\gamma}{1-\gamma+} [10 + \gamma V^*(S_0)] \end{aligned}$$

$$\begin{aligned} V^*(S_2) &= R(S_2) + max_a \gamma [T(S_2, a_0, S_0)V^*(S_0) + T(S_2, a_0, S_3)V^*(S_3)] \\ &= 1 + max_a \gamma [(1-y)V^*(S_0) + yV^*(S_3)] \\ &= 1 + \gamma [(1-y)V^*(S_0) + yV^*(S_3)] \\ &= 1 + 10y\gamma + (1 - y\gamma + y\gamma^2)V^*(S_0) \end{aligned}$$

$$
\begin{aligned}
V^*(S_3) &= R(S_3) + max_a\gamma[T(S_3, a_0, S_0)V^*(S_0)] \\
&= 10 + max_a\gamma[V^*(S_0)] \\
&= 10 + \gamma V^*(S_0)
\end{aligned}
$$

3 Q3

If there exist a x that makes

$$\pi(S_0) : S_0 \to a_2$$

so it means

$$V^*(S_2) \geq V^*(S_1)$$

Then we need to know

$$V^*(S_1) - V^*(S_2) = \frac{x\gamma}{1-\gamma+x\gamma}[10 + \gamma V^*(S_0)] - [1 + 10y\gamma + (1 - y\gamma + y\gamma^2)V^*(S_0)]$$

Considering

$$V^*(S_0) = 0$$

so we can calculate

$$
\begin{aligned}
V^*(S_1) - V^*(S_2) &= \frac{x\gamma}{1-\gamma+x\gamma} * 10 - (1 + 10y\gamma) \\
&= \frac{10x\gamma - (1+10y\gamma)(1-\gamma+x\gamma)}{1-\gamma+x\gamma} \\
&= \frac{(9-10y)x\gamma - (1+10y\gamma)(1-\gamma)}{1-\gamma+x\gamma}
\end{aligned}
$$

We need to discuss about different situation. Firstly, if y is greater than 0.9, so

$$9 - 10y \leq 0$$

there always exists

$$V^*(S_2) \geq V^*(S_1)$$

If y is less than 0.9, we need to know the minimal value of x. With the value of y increasing, the function "9-10y" is monotonically decreasing and the function "1+10yγ " is monotonically increasing. We observe that if we want to have the minimal value of x, we need to increase the value of the function

"9-10y" and decrease the value of the function "1+10yγ ". In that case, we choose

$$y = 0$$

Then we can simplify our calculation

$$V^*(S_1) - V^*(S_2) = \frac{9x\gamma - (1-\gamma)}{1-\gamma+x\gamma}$$

As

$$V^*(S_2) \geq V^*(S_1)$$

so

$$x \leq \frac{1-\gamma}{9\gamma}$$

With γ is greater than 0 but less than 1, we can know

$$x = 0$$

There exist a value of x (x = 0) that for all y $\in [0, 1] and \gamma \in (0, 1)$, $\pi(S_0) : S_0 \rightarrow a_2$

4 Q4

In the question Q3 we have

$$\begin{aligned} V^*(S_1) - V^*(S_2) &= \frac{x\gamma}{1-\gamma+x\gamma} * 10 - (1 + 10y\gamma) \\ &= \frac{10x\gamma - (1+10y\gamma)(1-\gamma+x\gamma)}{1-\gamma+x\gamma} \\ &= \frac{9x\gamma - 1 + \gamma - 10y\gamma(1-\gamma+x\gamma)}{1-\gamma+x\gamma} \end{aligned}$$

If we want $\pi(S_0) : S_0 \rightarrow a_1$ that means

$$\begin{aligned} V^*(S_2) &\leq V^*(S_1) \\ V^*(S_1) - V^*(S_2) &\geq 0 \\ \frac{9x\gamma - 1 + \gamma - 10y\gamma(1-\gamma+x\gamma)}{1-\gamma+x\gamma} &\geq 0 \end{aligned}$$

With $1 - \gamma + x\gamma$ is always greater than 0, we can simplify it:

$$\begin{aligned} 9x\gamma - 1 + \gamma - 10y\gamma(1-\gamma+x\gamma) &\geq 0 \\ \frac{9x\gamma - 1 + \gamma}{10\gamma(1-\gamma+x\gamma)} &\geq y \\ \frac{1}{10\gamma}(9 - \frac{10+8\gamma}{1-\gamma+x\gamma}) &\geq y \end{aligned}$$

With x $\in (0, 1)$,

$$9 - \frac{10 + 8\gamma}{1 - \gamma + x\gamma} \in (9 - \frac{10 + 8\gamma}{1 - \gamma}, -1)$$

That means the value of y is always less than 0, so there isn't any value of y that satisfy the requirement.

5 Q5

$$\pi^* = \{a1, a0, a0, a0\}$$

$$V^* = \{14.18521119, 15.76139333, 15.69747038, 22.76664727\}$$

python code: https://github.com/SummerOf15/machine-learning-in-robots