Lab2 - Q-Learning

Maxime Lhoustau and Philipp Schlieker

September 29, 2019

Question 1

1. $\pi(S_0) = a_1$
2. $\pi(S_0) = a_2$
3. $\pi(S_1) = a_0$
4. $\pi(S_2) = a_0$
5. $\pi(S_3) = a_0$

Question 2

- $V^*(S_0) = R(S_0) + max_a \gamma \sum_{S'} T(S_0, a, S')V^*(S') =$
  $0 + max(\gamma T(S_0, a_1, S_1)V^*(S_1), \gamma T(S_0, a_2, S_2)V^*(S_2)) =$
  $max(\gamma V^*(S_1), \gamma V^*(S_2))$
- $V^*(S_1) = R(S_1) + max_a \gamma \sum_{S'} T(S_1, a, S')V^*(S') =$
  $0 + max(\gamma(T(S_1, a_0, S_1)V^*(S_1) + T(S_1, a_0, S_3)V^*(S_3)) =$
  $\gamma((1 - x)V^*(S_1) + xV^*(S_3))$
- $V^*(S_2) = R(S_2) + max_a \gamma \sum_{S'} T(S_2, a, S')V^*(S') =$
  $1 + max(\gamma(T(S_2, a_0, S_0)V^*(S_0) + T(S_2, a_0, S_3)V^*(S_3))) =$
  $1 + \gamma((1 - y)V^*(S_0) + yV^*(S_3))$
- $V^*(S_3) = R(S_3) + max_a \gamma \sum_{S'} T(S_3, a, S')V^*(S') =$
  $10 + max(\gamma T(S_3, a_0, S_0)V^*(S_0)) =$
  $10 + \gamma S_0$

Question 3

Let $x = 0$ then $V^*(S_1) = 0 + \gamma V^*(S_1) = 0 \Rightarrow \forall y \in [0,1] : V^*(S_1) < V^*(S_2) \Rightarrow \pi^*(S_0) = a_1$

Question 4

$\nexists y \in [0,1]$ s.t. $\forall\ V \in [0,1] : \pi^*(S_0) = a_0$.
Proof by contradiction: If such y exists for which $V^*(S_1) > V^*(S_2)$, then $V^*(S_1) > V^*(S_2)$ must also hold for the y that minimizes $V^*(S_2)$.
$arg_y min V^*(S_2) = 0$. Let $y = 0, x = 0.00001$ then $V^*(S_1) < V^*(S_2) \Rightarrow \pi^*(S_0) = a_2$
Hence no such y can exist.

Question 5

$V = [14.18, 15.76, 15.70, 22.76]$
$\pi = [1, 0, 0, 0]$