Assignment 5

Reinforcement Learning: assignment #1

Davide Brinati

Matricola: 771458

a) Define the MDP underling the problem.

- S = {s1, s2, s3, s4, s5, s6, s7}
- A = {Destra, Sinistra}
- T = Matrice di transizione per un processo deterministico, T(s,a) → s'
  - T(s1, Destra) = s3
  - T(s1, Sinistra) = s2
  - T(s2, Destra) = s5
  - T(s2, Sinistra) = s4
  - T(s3, Destra) = s7
  - T(s3, Sinistra) = s6
- R = funzione reward o payoff, R(s)
  - R(s1, Destra) = 0
  - R(s1, Sinistra) = 0
  - R(s2, Destra) = 0
  - R(s2, Sinistra) = 4
  - R(s3, Destra) = 3
  - R(s3, Sinistra) = 2

b) Compute the optimal value function V* using the value iteration algorithm.

Assumiamo che $\gamma = 1$, utilizzando l'algoritmo di iterazione del valore avremo che:

- V(s4) = 0
- V(s5) = 0
- V(s6) = 0
- V(s7) = 0

- V(s2) = max{4, 0} = 4
- V(s3) = max{2, 3} = 3
- V(s1) = max{0 + 4, 0 + 3} = 4

c) Suppose to have an initial policy which chooses equally between right and left at each junction, and assume γ = 0.5:

1. What is the value function V^{π}_0 for the initial policy?

La *initial policy* ci da i seguenti risultati:

- A s1 sceglie Destra
- A s2 sceglie Destra
- A s3 sceglie Sinistra

Quindi il valore della funzione V^{π}_0 sarà:

- $V^{\pi}_0(s1)$ = max{0 + 0.5*0; 0 + 0.5*2} = max{0, 1} = 1 (va a destra)
- $V^{\pi}_0(s2)$ = max{4, 0} = 4 (va a sinistra)
- $V^{\pi}_0(s3)$ = max{2, 3} = 3 (va a destra)

2. What is the improved policy π1 based on the value function V^{π}_0? [Policy Improvement]

- A s1 sceglie di andare a Destra
- A s2 sceglie di andare a Sinistra
- A s3 sceglie di andare a Destra

3. What is the new value function V^{π}_1? [Policy Evaluation]

- $V^{\pi}_1(s1)$ = max{0 + 0.5*4; 0 + 0.5*3} = max{2, 1.5} = 2 (va a sinistra)
- $V^{\pi}_1(s2)$ = max{4, 2} = 4(va a sinistra)
- $V^{\pi}_1(s3)$ = max{2, 3} = 3 (va a destra)

4. **What is the improved policy $\pi 2$ based on the value function $V\pi 1$? [Policy Improvement]**

- A s1 sceglie di andare a sinistra
- A s2 sceglie di andare a sinistra
- A s3 sceglie di andare a destra