



# Q\_Learning

# Q-Learning - Exercício

- Considere o ambiente 3 x 2 onde a posição 6 é o estado terminal
- Assuma que as ações UP, DOWN, LEFT e RIGHT são determinísticas
- Recompensas:
  - +10 no estado 6
  - -10 se bater na parede
  - -1 nos outros casos
- Aplicar o Q-learning sequencialmente usando as seguintes trajetórias:
  - Estado inicial 1, sequência U,U,U,R
  - Estado inicial 5: sequência L,L,U,R
- Inicialize a matriz Q com zeros e assumo  $\alpha = 0.5$  e  $\gamma = 0.8$

3	6
2	5
1	4