

Inteligência Artificial

Flávio Soares Corrêa da Silva

1º. Semestre 2021



Exercícios

- Nesta segunda lista exploraremos o algoritmo Q-learning. Para isso, teremos um problema de navegação com informações conhecidas por nós, mas que precisarão ser descobertas pelo agente. O objetivo será simular o Q-learning para o agente conseguir identificar uma política “boa”. Consideraremos duas variações de um problema com 12 estados, apenas um estado terminal e ações com confiabilidade exatamente como no exercício visto nas aulas anteriores. As recompensas e confiabilidade das ações serão desconhecidas inicialmente pelo agente.

Exercícios

- *In this second list, we will explore the Q-learning algorithm. We will have a navigation problem about which we will know every relevant information, some of which will need to be discovered by the agent. The goal will be to simulate Q-learning so that the agent can identify a “good” policy. We will consider two variations of a problem containing 12 states, only one terminal state and reliability of actions exactly as in the example seen during lectures. Rewards and reliability of actions shall be unknown, initially, by the agent.*

Exercício 1

recompensas

Estado inicial

-1	-1	-1	+100
-1	-100	-1	-1
-1	-1	-1	-1

Estado termina

Exercício 1

- O resultado do Q-learning corresponde ao esperado intuitivamente?
 - Quantas rodadas foram necessárias para chegar a este resultado?
-
- *Does the result provided by Q-learning correspond to what would be expected intuitively?*
 - *How many rounds were required to reach this result?*

Exercício 2

recompensas

Estado inicial

-1	-1	-1	+100
-1	-100	-1	-1
-1	-1	-1	+10

Estado termina

Recompensa modificada

Exercício 2

- O resultado do Q-learning corresponde ao esperado intuitivamente?
 - Em caso negativo, é possível ajustar as informações dos estados para induzir o Q-learning a produzir o resultado desejado?
-
- *Does the result provided by Q-learning correspond to what would be expected intuitively?*
 - *In case it does not, is it possible to adjust the informations about states in order to induce Q-learning to produce the desired result?*