

## Inteligência Artificial

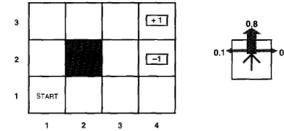
Exercícios e Respostas – Cap. 17

Profª Bianca Zadrozny

<http://www.ic.uff.br/~bianca/ia>

## Exercícios – Cap. 17

1. (17.1) Para o grid abaixo, indique quais quadrados podem ser atingidos pela sequência de ações [Cima,Cima,Direita, Direita, Direita] e com quais probabilidades.



## Exercícios – Cap. 17

2. (17.4) Considere um PDM (sem descontos) com três estados (1, 2, 3), com recompensas -1, -2, 0, respectivamente. O estado 3 é terminal. Nos estados 1 e 2 há duas ações possíveis: *a* e *b*. O modelo de transição é dado a seguir:
- No estado 1, a ação *a* move o agente para o estado 2 com prob. 0,8 e faz ele ficar no mesmo lugar com prob. 0,2.
  - No estado 2, a ação *a* move o agente para o estado 1 com prob. 0,8 e faz ele ficar no mesmo lugar com prob. 0,2.
  - Nos estados 1 e 2, a ação *b* move o agente para o estado 3 com prob. 0,1 e faz ele ficar no mesmo lugar com prob. 0,9.
- a) O que você pode dizer qualitativamente sobre a política ótima para os estados 1 e 2?
- b) Aplique iteração de valor para achar a política ótima para os estados 1 e 2.

## Respostas – Cap. 17

1. (3,1): 0,01  
 (3,2): 0,08  
 (3,3): 0,09  
 (4,2): 0,18  
 (4,3): 0,64

## Respostas – Cap. 17

2. a) O agente deve tentar chegar logo ao estado terminal 3 para parar de receber recompensas negativas. Porém, a única ação que leva ao estado 3 (*b*) tem baixa chance de sucesso; então o agente deve tentar minimizar o custo de tentar chegar a 3. Então, a partir de 1, o agente deve tentar chegar a 3 (executando *b*). Mas a partir de 2, o agente deve primeiro executar *a* pra chegar a 1 e só a partir daí começar a tentar chegar em 3 a partir de um estado em que ele perde menos.

## Respostas – Cap. 17

2. b) 
$$U_{i+1}(s) \leftarrow R(s) + \gamma \max_a \sum_{s'} T(s, a, s') U_i(s')$$
- $U_0(1) = 0; U_0(2) = 0; U_0(3) = 0;$   
 $U_1(1) = -1; U_1(2) = -2; U_1(3) = 0;$   
 $U_2(1) = -1 + \max(0.8*(-2) + 0.2*(-1), 0.1*0 + 0.9*(-1))$   
 $= -1 - 0.9 = -1.9$   
 $U_2(2) = -2 + \max(0.8*(-1) + 0.2*(-2), 0.1*0 + 0.9*(-2))$   
 $= -2 - 1.2 = -3.2$   
 $U_2(3) = 0$

## Respostas – Cap. 17

$$U_3(1) = -1 + \max(0.8*(-3.2) + 0.2*(-1.9), 0.1*0 + 0.9*(-1.9)) \\ = -1 - 1.7 = -2.7$$

$$U_3(2) = -2 + \max(0.8*(-1.9) + 0.2*(-3.2), 0.1*0 + 0.9*(-3.2)) \\ = -2 - 2.9 = -4.9$$

$$U_3(3) = 0$$

$$U_4(1) = -1 + \max(0.8*(-4.9) + 0.2*(-2.7), 0.1*0 + 0.9*(-2.7)) \\ = -1 - 2.4 = -3.4$$

$$U_4(2) = -2 + \max(0.8*(-2.7) + 0.2*(-4.9), 0.1*0 + 0.9*(-4.9)) \\ = -2 - 4.4 = -6.4$$

$$U_4(3) = 0$$

...