

Lista 3 - MAC0425

Lucas Quaresma Medina Lam

Nº USP: 11796399

1-

Ações em STRIPS:

Ação: Ir(x, y, r)

Pré-cond: Sobre(Shakey, Piso) \wedge Em(Shakey, x) \wedge Em(x, r) \wedge Em(y, r) \wedge Sala(r)

Efeito: Em(Shakey, y) \wedge \sim Em(Shakey, x)

Ação: Empurrar(b, x, y, r)

Pré-cond: Sobre(Shakey, Piso) \wedge Em(Shakey, x) \wedge Em(b, x) \wedge Em(x, r) \wedge Em(y, r) \wedge Sala(r) \wedge Caixa(b)

Efeitos: Em(Shakey, y) \wedge Em(b, y) \wedge \sim Em(Shakey, x) \wedge \sim Em(b, x)

Ação: Subir(x, b)

Pré-cond: Em(Shakey, x) \wedge Em(b, x) \wedge Sobre(Shakey, Piso)

Efeitos: Sobre(Shakey, b) \wedge \sim Sobre(Shakey, Piso)

Ação: Descer(b, x)

Pré-cond: Em(Shakey, x) \wedge Em(b, x) \wedge Sobre(Shakey, b)

Efeitos: Sobre(Shakey, Piso) \wedge \sim Sobre(Shakey, b)

Ação: Ligar(s, b)

Pré-cond: $\text{Sobre}(\text{Shakey}, b) \wedge \text{Em}(b, x) \wedge \text{Em}(s, x) \wedge \text{Posicao}(x) \wedge \text{Interruptor}(s) \wedge \text{Desligado}(s)$

Efeitos: $\text{Ligado}(s) \wedge \sim \text{Desligado}(s)$

Ação: $\text{Desligar}(s, b)$

Pré-cond: $\text{Sobre}(\text{Shakey}, b) \wedge \text{Em}(b, x) \wedge \text{Em}(s, x) \wedge \text{Posicao}(x) \wedge \text{Interruptor}(s) \wedge \text{Ligado}(s)$

Efeitos: $\text{Ligado}(s) \wedge \sim \text{Desligado}(s)$

Estado Inicial:

$\text{Sala}(\text{SALA1}), \text{Sala}(\text{SALA2}), \text{Sala}(\text{SALA3}), \text{Sala}(\text{SALA4}), \text{Sala}(\text{CORREDOR}),$
 $\text{Em}(\text{PORTA1}, \text{SALA1}), \text{Em}(\text{PORTA1}, \text{CORREDOR}),$
 $\text{Em}(\text{PORTA2}, \text{SALA2}), \text{Em}(\text{PORTA2}, \text{CORREDOR}), \text{Em}(\text{PORTA3}, \text{SALA3}),$
 $\text{Em}(\text{PORTA3}, \text{CORREDOR}), \text{Em}(\text{PORTA4}, \text{SALA4}),$
 $\text{Em}(\text{PORTA4}, \text{CORREDOR}), \text{Caixa}(\text{CAIXA1}), \text{Caixa}(\text{CAIXA2}), \text{Caixa}(\text{CAIXA3}),$
 $\text{Caixa}(\text{CAIXA4}), \text{Em}(\text{Shakey}, \text{SALA3}), \text{Em}(\text{CAIXA1}, \text{SALA1}),$
 $\text{Em}(\text{CAIXA2}, \text{SALA1}), \text{Em}(\text{CAIXA3}, \text{SALA1}), \text{Em}(\text{CAIXA4}, \text{SALA1}), \text{Em}(a, \text{SALA2}),$
 $\text{Em}(\text{Caixa2}, b), \text{Em}(b, \text{SALA1}), \text{Em}(\text{Shakey}, c), \text{Em}(c, \text{SALA3})$

Plano para Shakey colocar caixa2 em sala2:



2-a)

Preenchendo as linhas que falta, temos:

Discount Factor, $\gamma = 0.5$

| s | a | s' | T(s,a,s') | R(s,a,s') |
|---|------------------|----|-----------|-----------|
| A | Clockwise | B | 0.6 | 0 |
| A | Clockwise | C | 0.4 | 1 |
| A | Counterclockwise | B | 0.600 | 0.000 |
| A | Counterclockwise | C | 0.400 | 9.000 |
| B | Clockwise | C | 1.000 | 0.000 |
| B | Counterclockwise | A | 1.000 | -2.000 |
| C | Clockwise | A | 0.800 | 0.000 |
| C | Clockwise | B | 0.200 | -6.000 |
| C | Counterclockwise | B | 1.000 | 6.000 |

b)

$$Q_{k+1}(A, \text{Clockwise}) = (((0.6)[0 + 0.5 * 3]) + ((0.4)[1 + 0.5 * 6])) = 2.5$$

$$Q_{k+1}(B, \text{Clockwise}) = (((0) + ((1)[0 + 0.5 * 6])) = 3$$

$$Q_{k+1}(C, \text{Clockwise}) = (((0.8)[0 + 0.5 * 4.8]) + ((0.2)[-6 + 0.5 * 3])) = 1.02$$

$$Q_{k+1}(A, \text{Counterclockwise}) = (((0.6)[0 + 0.5 * 3]) + ((0.4)[9 + 0.5 * 6])) = 5.7$$

$$Q_{k+1}(B, \text{Counterclockwise}) = (((1)[-2 + 0.5 * 4.8]) + ((0))) = 0.4$$

$$Q_{k+1}(C, \text{Counterclockwise}) = (((0) + ((1)[6 + 0.5 * 3])) = 7.5$$

3-a)

Construindo os Q valores para os 5 episódios, temos:

Episódio 0:

Inicializa todos os Q valores com 0.

$Q_0((m,n), a) = 0$. Sendo 'm' e 'n' as posições e 'a' a ação

Episódio 1:

$$Q((1, 3), S) = 0$$

$$Q((1, 2), E) = 0$$

$$Q((2, 2), E) = 0$$

$$Q((3, 2), N) = 0$$

$$Q((3, 3), \text{Exit}) = 0.5 \cdot 0 + 0.5 \cdot (50 + 0.5 \cdot 0) = 25$$

Episódio 2:

$$Q((1, 3), S) = 0$$

$$Q((1, 2), E) = 0$$

$$Q((2, 2), S) = 0$$

$$Q((2, 1), \text{Exit}) = 0.5 \cdot 0 + 0.5 \cdot (-100 + 0.5 \cdot 0) = -50$$

Episódio 3:

$$Q((1, 3), S) = 0$$

$$Q((1, 2), E) = 0$$

$$Q((2, 2), E) = 0$$

$$Q((3, 2), S) = 0$$

$$Q((3, 1), \text{Exit}) = 30 \cdot 0.5 + 0 = 15$$

Episódio 4:

$$Q((1, 3), S) = 0$$

$$Q((1, 2), E) = 0$$

$$Q((2, 2), E) = 0$$

$$Q((3, 2), N) = 0.5 \cdot 0 + 0.5 \cdot (0 + 0.5 \cdot 25) = 6.25$$

$$Q((3, 3), Exit) = 0.5 \cdot 25 + 0.5 \cdot (50 + 0.5 \cdot 0) = 37.5$$

Episódio 5:

$$Q((1, 3), S) = 0$$

$$Q((1, 2), E) = 0$$

$$Q((2, 2), E) = 0.5 \cdot 0 + 0.5 \cdot (0 + 0.5 \cdot 6.25) = 1.5625$$

$$Q((3, 2), S) = 0.5 \cdot 0 + 0.5 \cdot (0 + 0.5 \cdot 15) = 3.75$$

$$Q((3, 1), Exit) = 0.5 \cdot 15 + 0.5 \cdot (30 + 0.5 \cdot 0) = 22.5$$

Então, para os valores pedidos no enunciado, temos:

$$Q((3, 2), N) = 0.5 \cdot 0 + 0.5 \cdot (0 + 0.5 \cdot 25) = 6.25$$

$$Q((3, 2), S) = 0.5 \cdot 0 + 0.5 \cdot (0 + 0.5 \cdot 15) = 3.75$$

$$Q((2, 2), E) = 0.5 \cdot 0 + 0.5 \cdot (0 + 0.5 \cdot 6.25) = 1.5625$$

b)

$$w_i \leftarrow w_i + \alpha [r + \gamma [\max_{a'} Q_f(s', a')] - Q_f(s, a)] f_i(s, a)$$

$$(i) Q_f(s, a) = w_1 \cdot f_1(s) + w_2 \cdot f_2(s) + w_3 \cdot f_3(a)$$

No primeiro episódio, só atualizaríamos os valores dos pesos em ações que tenham recompensa, então temos que os valores dos pesos seriam:

$$diff = 50 + 0.5 \cdot 0 - 0 = 50$$

$$w_1 = 0 + 0.5 \cdot 50 \cdot 3 = 75$$

$$w_2 = 0 + 0.5 \cdot 50 \cdot 3 = 75$$

$$w_3 = 0 + 0.5 \cdot 50 \cdot 1 = 25$$

(ii)

Ao variar as ações, vemos que os valores de f_1 e f_2 não se alteram. Portanto, a melhor ação recomendada pela função Q no estado (2,2) seria aquela que maximizasse o valor f_3 . Ou seja, a ação W seria a recomendada, pois tem valor $f_3 = 4$.