Lista 3 - MAC0425

Lucas Quaresma Medina Lam

N° 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) \land \simEm(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) \land Em(b, y) \land~Em(Shakey, x) \land~Em(b, x)
Ação: Subir(x, b)
Pré-cond: Em(Shakey, x) \land Em(b, x) \land Sobre(Shakey, Piso)
Efeitos: Sobre(Shakey, b) ∧ ~Sobre(Shakey, Piso)
Ação: Descer(b, x)
Pré-cond: Em(Shakey, x) \land Em(b, x) \land Sobre(Shakey, b)
Efeitos: Sobre(Shakey, Piso) \land \simSobre(Shakey, b)
Ação: Ligar(s, b)
```

Lista 3 - MAC0425

Pré-cond: Sobre(Shakey, b) \land Em(b, x) \land Em(s, x) \land Posicao(x) \land Interruptor(s) \land Desligado(s)

Efeitos: Ligado(s) $\land \sim$ Desligado(s)

Ação: Desligar(s, b)

Pré-cond: Sobre(Shakey, b) \land Em(b, x) \land Em(s, x) \land Posicao(x) \land Interruptor(s) \land

Ligado(s)

Efeitos: Ligado(s) $\land \sim$ Desligado(s)

Estado Inicial:

Sala(SALA1), Sala(SALA2), Sala(SALA3), Sala(SALA4), Sala(CORREDOR),

Em(PORTA1, SALA1), Em(PORTA1, CORREDOR),

Em(PORTA2, SALA2), Em(PORTA2, CORREDOR), Em(PORTA3, SALA3),

Em(PORTA3, CORREDOR), Em(PORTA4, SALA4),

Em(PORTA4, CORREDOR), Caixa(CAIXA1), Caixa(CAIXA2), Caixa(CAIXA3),

Caixa(CAIXA4), Em(Shakey, SALA3), Em(CAIXA1, SALA1),

Em(CAIXA2, SALA1), Em(CAIXA3, SALA1), Em(CAIXA4, SALA1), Em(a, SALA2),

Em(Caixa2, b), Em(b, SALA1), Em(Shakey, c), Em(c, SALA3)

Plano para Shakey colocar caixa2 em sala2:

Lista 3 - MAC0425 2

INIT Sala(SALA1), Sala(SALA2), Sala(SALA3), Sala(SALA4), Sala(CORREDOR), Em(PORTA1, SALA1), Em(PORTA1, CORREDOR), Em(PORTA2, SALA2), Em(PORTA2, CORREDOR), Em(PORTA3, SALA3), Em(PORTA3, CORREDOR), Em(PORTA4, SALA4), EM(PORT 44, CORREDOR), Caixa(CAIXA1), Caixa(CAIXA2), Caixa(CAIXA3), Caixa(CAIXA4), EM(CAIXA1, SALA1), EM (CAIXA2, SALA1), EM(CAIXA3, SALA1), EM(CAIXA4, SALA1), EM(a, SALA2), EM(Caixa2, b), EM(b, SALA1), EM(Shakey, c), EM(c, SALA3) Init Init Init Sobre(Shakey, Piso), Em(Shakey, c), Em(PORTA3, SALA3, Em(c, SALA3), Sala(SALA3) Ir(c, PORTA3, SALA3) Em(Shakey, PORTA3), ~Em(Shakey, c) Init Sobre(Shakey, Piso), Em(Shakey, PORTA3), Em(PORTA1, CORREDOR), Em(PORTA3, CORREDOR), Sala(CORREDOR)

In(PORTA3, PORTA1, CORREDOR)

Em(Shakey, PORTA1), ~Em(Shakey, PORTA3) Init Sobre(Shakey, Piso), Em(Shakey, PORTA1), Em(b, SALA1), Em(PORTA1, SALA1), Sala(SALA1) Ir(PORTA1, b, SALA1) Em(Shakey, b), ~Em(Shakey, PORTA1) Init Init Init Init Sobre(Shakey, Piso), Em(Shakey, b), Em(CAIXA2, b), Em(b, SALA1), Em(PORTA1, SALA1), Sala(SALA1), Caixa(CAIXA2) Empurrar(CAIXA2, 6, PORTA1, SALA1) Em(Shakey, PORTA1), Em(CAIXA2, PORTA1), ~Em(Shakey, b), ~Em(CAIXA2, b) Init Init Sobre(Shakey, Piso), Em(Shakey, PORTAI), Em(CAIXA2, PORTAI), Em(PORTAI, CORREDOR), Em(PORTA2, CORREDOR), Sala(CORREDOR), Caixa(CAIXA2) Empurrar(CAIXA2, PORTA1, PORTA2, CORREDOR) Em(Shakey, PORTA2), Em(CAIXA2, PORTA2), ~Em(Shakey, PORTA1), ~Em(CAIXA2, PORTA1) Init Init Sobre(Shakey, Piso), Em(Shakey, PORTA2), Em(CAIXA2, PORTA2), Em(PORTA2, SALA2), Em(a, SALA2), Sala(SALA2), Caixa(CAIXA2) Empurrar(CAIXA2, PORTA2, a, SALA2) Em(Shakey, a), Em(CAIXA2, a), ~Em(Shakey, PORTA2), ~Em(CAIXA2, PORTA2) Em(CAIXA2, a), Em(a, SALA2) GOAL

2-a)

Preenchendo as linhas que falta, temos:

Lista 3 - MAC0425 3

Discount Factor, $\gamma = 0.5$

s	а	s'	T(s,a,s')	R(s,a,s')
Α	Clockwise	В	0.6	0
Α	Clockwise	С	0.4	1
Α	Counterclockwise	В	0.600	0.000
Α	Counterclockwise	С	0.400	9.000
В	Clockwise	С	1.000	0.000
В	Counterclockwise	Α	1.000	-2.000
С	Clockwise	Α	0.800	0.000
С	Clockwise	В	0.200	-6.000
С	Counterclockwise	В	1.000	6.000

b)

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

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

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

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

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

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

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:

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

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

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

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

$$Q((3, 3), Exit) = 0.5*0 + 0.5*(50 + 0.5*0) = 25$$

Episódio 2:

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

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

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

$$Q((2, 1), Exit) = 0.5*0 + 0.5*(-100 + 0.5*0) = -50$$

Episódio 3:

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

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

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

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

$$Q((3, 1), Exit) = 30 * 0.5 + 0 = 15$$

Episódio 4:

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

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

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

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

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

Episódio 5:

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

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

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

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

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

Então, para os valores pedidos no enunciado, temos:

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

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

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

b)

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

(i)
$$Qf(s,a) = w1.f1(s) + w2.f2(s) + w3.f3(a)$$

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

$$diff = 50 + 0.5*0 - 0 = 50$$

$$w_1 = 0 + 0.5 * 50 * 3 = 75$$

$$w_2 = 0 + 0.5 * 50 * 3 = 75$$

$$w_3 = 0 + 0.5 * 50 * 1 = 25$$

(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 $f_3=4$.

Lista 3 - MAC0425 7