Value Interation (MDP)
Inteligênica Artificial – PCS3438

Arthur Font – 12036152

Iteration 0

0	0	0	0
0	0	0	0
		0	0
(D) 0	0	0	0

V(s) = 0, for all s

Iteration 1

-1	-1	-1	-1
-1	-1	-1	-1
		-1	-1
(D) 0	-1	-1	-1

V(s) = max a ((r(s, O) + V(s' 1)),
(r(s, N) + V(s' 2)),
(r(s, L) + V(s' 3)),
(r(s, S) + V(s' 4)))
= max a ((-1+0), (-1+0), (-1+0), (-1+0))
= -1

Iteration 2

-2	-2	-2	-2
-2	-2	-2	-2
		-2	-2
(D) 0	-1	-2	-2

Iteration 3

-3	-3	-3	-3
-3	-3	-3	-3
		-3	-3
(D) 0	-1	-2	-3

Iteration 4

-4	-4	-4	-4
-4	-4	-4	-4
		-3	-4
(D) 0	-1	-2	-3

Iteration 5

-5	-5	-5	-5
-5	-5	-4	-5
		-3	-4
(D) 0	-1	-2	-3

Iteration 6

-6	-6	-5	-6
-6	-5	-4	-5
		-3	-4
(D) 0	-1	-2	-3

Iteration 7

-7	-6	-5	-6
-6	-5	-4	-5
		-3	-4
(D) 0	-1	-2	-3

Iteration 8

-7	-6	-5	-6
-6	-5	-4	-5
		-3	-4
(D) 0	-1	-2	-3

$V(s) - V(s') = 0$, for all s → Convergência do VI

1) Qual estado tem valor mínimo? Qual o valor deste estado?
R: S(1,4). O valor deste estado é -7.

2) Qual estado tem valor máximo? Qual o valor deste estado?
R: S(1,1). O valor deste estado é 0.

3) Mostrar na grade qual é a política ótima em cada célula.
R:

-7 ↓→	-6 ↓→	-5 ↓	←↓ -6
-6 →	-5 →	-4 ↓	←↓ -5
		-3 ↓	←↓ -4
(D) 0	← -1	← -2	← -3