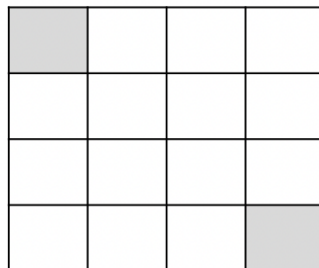Haonan Hu

4/27/2021

2863545

In-Class Problem

1. Use the Value Iteration algorithm to find the Optimal Policy (π*) for the Gridworld task below with the two termination states shown as gray squares below. Make all the assumptions I did in the example. Show your calculations.
2. How many iterations did it take to converge?

K=0

0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

K=1

0	-1	-1	-1
-1	-1	-1	-1
-1	-1	-1	-1
-1	-1	-1	0

V(s) = max[-1+1(0), -1+1(0), -1+1(0), -1+1(0)] = max[-1,-1,-1,-1] = -1

K=2

0	-1	-2	-2
-1	-2	-2	-2
-2	-2	-2	-1
-2	-2	-1	0

V(s) = max[-1+1(0), -1+1(-1), -1+1(-1), -1+1(-1)] = max[-1,-2,-2,-2] = -1

V(s) = max[-1+1(-1), -1+1(-1), -1+1(-1), -1+1(-1)] = max[-2,-2,-2,-2] = -2

K=3

0	-1	-2	-3
-1	-2	-3	-2
-2	-3	-2	-1
-3	-2	-1	0

V(s) = max[-1+1(0), -1+1(-1), -1+1(-2), -1+1(2)] = max[-1,-2,-3,-3] = -1

V(s) = max[-1+1(-1), -1+1(-2), -1+1(-2), -1+1(-2)] = max[-2,-3,-3,-2] = -2

V(s) = max[-1+1(-2), -1+1(-2), -1+1(-2), -1+1(-2)] = max[-3,-3,-3,-3] = -3

K=4

0	-1	-2	-3
-1	-2	-3	-2
-2	-3	-2	-1
-3	-2	-1	0

V(s) = max[-1+1(0), -1+1(-1), -1+1(-2), -1+1(2)] = max[-1,-2,-3,-3] = -1

V(s) = max[-1+1(-1), -1+1(-2), -1+1(-2), -1+1(-2)] = max[-2,-3,-3,-2] = -2

V(s) = max[-1+1(-2), -1+1(-2), -1+1(-2), -1+1(-2)] = max[-3,-3,-3,-3] = -3

2. How many iterations did it take to converge?

4 iterations