

TEAM 18

Vidit Jain (201501021)

Harshit Patni (201501107)

Value Iteration :

Matrices Till Convergence :

Initial State :

W	W	18	W
0	0	0	0
0	-18	W	0
0	0	0	0

After 1 Iterations

W	W	18	W
-0.9000	-0.9000	13.5000	-0.9000
-0.9000	-18	W	-0.9000
-0.9000	-0.9000	-0.9000	-0.9000

After 2 Iterations

W	W	18	W
-1.8000	8.0100	13.3200	9.7200
-1.8000	-18	W	-1.8000
-1.8000	-1.8000	-1.8000	-1.8000

After 3 Iterations

W	W	18	W
5.1480	8.7570	15.2730	10.5480
-2.7000	-18	W	6.5160
-2.7000	-2.7000	-2.7000	-2.7000

After 4 Iterations

W	W	18	W
6.3504	10.3941	15.4305	13.0248
1.1484	-18	W	8.8416
-3.6000	-3.6000	-3.6000	3.7728

After 5 Iterations

W	W	18	W
8.1652	10.6838	15.8419	13.6310
2.4952	-18	W	11.2882
-0.7013	-4.5000	1.3982	6.1906

After 6 Iterations

W	W	18	W
8.7131	11.0419	15.9315	14.2654
4.0816	-18	W	12.2625
0.5760	-2.0314	4.3321	8.8894

After 7 Iterations

W	W	18	W
9.2130	11.1494	16.0307	14.4980
4.6786	-18	W	12.9648
2.2198	0.5625	7.0779	10.2321

After 8 Iterations

W	W	18	W
9.4087	11.2395	16.0647	14.6709
5.1383	-18	W	13.2913
3.1211	3.0186	8.7013	11.2029

After 9 Iterations

W	W	18	W
9.5463	11.2757	16.0910	14.7480
5.3408	-18	W	13.4950
3.8246	4.5629	9.8026	11.7235

After 10 Iterations

W	W	18	W
9.6093	11.3004	16.1024	14.7971
5.4711	-18	W	13.5974
4.2114	5.5983	10.4393	12.0486

After 11 Iterations

W	W	18	W
9.6484	11.3119	16.1098	14.8214
5.5346	-18	W	13.6572
4.5469	6.2113	10.8267	12.2267

After Convergence :

W	W	18	W
9.6484	11.3119	16.1098	14.8214
5.5346	-18	W	13.6572
4.5469	6.2113	10.8267	12.2267

Expected Reward :

The final Expected Reward : 4.5469

Optimal Path From Start to End :

Current State : 3 0

Action To Take : East

Current State : 3 1

Action To Take : East

Current State : 3 2

Action To Take : East

Current State : 3 3

Action To Take : North

Current State : 2 0

Action To Take : North

Current State : 2 3

Action To Take : North

Current State : 1 0

Action To Take : East

Current State : 1 1

Action To Take : East

Current State : 1 2

Action To Take : North

Current State : 1 3

Action To Take : West

Linear Programming :

Values of x:

State, Action pair	Value of X
3,5	0.864702570838806
5,1	0
5,2	0
5,3	0
5,4	0.121765601217656
6,1	0
6,2	0
6,3	0
6,4	0.228333669254418
7,1	1.08087821354851
7,2	0
7,3	0
7,4	0
8,1	0
8,2	0
8,3	1.12276409768122
8,4	0

9,1	0.136986301369863
9,2	0
9,3	0

9,4	0
10,5	0.135297429161194
12,1	1.12799983319781
12,2	0
12,3	0
12,4	0
13,1	0
13,2	0
13,3	0
13,4	1.11111111111111
14,1	0
14,2	0
14,3	0
14,4	0.987654320987654
15,1	0
15,2	0
15,3	0
15,4	1.11111111111111
16,1	0.987654320987655

16,2	0
16,3	0
16,4	0

Expected Reward:

5.914659828

Description of why the records match/don't match:

Our aim is to maximize utility/reward. Hence, both value iteration and linear programming, will end up achieving the same results if the precision is high and accurate i.e. $\delta = 0$. Larger deltas will provide less iterations and gives the approximate value of utilities at all the stages.

Verification of result:

Value using LP / Value using VI = $5.914659828 / 4.5469 = 1.3008115041$