Assignment #2

Sudhanshu Pathak

Q1:

NOTE: I have used Java for programming. Please run program in following format:

```
java MDPValueIteration filePath horizon
```

Results: Horizon = 10

Input:

3 2

0.2 0.8 0.0
0.0 0.2 0.8
1.0 0.0 0.0

0.9 0.05 0.05
0.05 0.9 0.05
0.05 0.05 0.9

-1.0 -1.0 0.0

Output:

```
Value Function Matrix
------------------------------------------------------------------
------
[-1.00000 -1.95000 -2.35000 -2.52600 -2.67480 -2.82888 -2.98618 -
3.14412 -3.30207 -3.45997 ]

[-1.00000 -1.20000 -1.32000 -1.46200 -1.61740 -1.77550 -1.93361 -
2.09155 -2.24945 -2.40734 ]

[0.00000 -0.10000 -0.24750 -0.40625 -0.56503 -0.72313 -0.88104 -
1.03892 -1.19681 -1.35471 ]

Policy Matrix
------------------------------------------------------------------
------
[0 1 0 0 0 0 0 0 0 0 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 1 1 1 1 1 1 1 1 1 ]
```

Q2: Reach Home MDP:

It's a simple MDP example in which the agent tries to reach home from school and on the way due to distractions it takes diversions and takes more time to reach home. For distractions there is a negative reward and for taking BUS it gets reward such as +6. The states indicated in green indicates positive reward state and states indicated in red states negative reward states. Following is an example of 5x4 = 20 states MDP.

Transition matrix:

	0	1	2	3
0	BUS	s4	s8	HOME
1	s3	s5	FIGHT	s12
2	s2	s6	s9	s13
3	s1	PUB	s10	s14
4	s0	s7	s11	s15

Actions:
Since it's a grid example, possible actions are Left, Right, Up and Down.

Action
Left
Right
Up
Down

State	Reward
BUS	6
PUB	-2
FIGHT	-4
HOME	10
Other states	-1

Transition Probability: Here are some of the rules for transition function.

1. For normal action agent reaches to ay S state, it gets reward of -1.
2. For special states like BUS, PUB, FIGHT, HOME, agent gets special reward as mentioned in the table above.

Results:

Input:

```
20 4

0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0.7  0    0    0    0.3  0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0.6  0    0.3  0    0    0    0.1  0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0.6  0    0    0    0    0
     0.3  0    0    0    0.1  0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    1    0    0    0
0    0    0    0    0    0    0.6  0    0    0    0    0
     0.3  0    0    0    0.1  0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0.1  0
     0    0    0.7  0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0
     0    0    0    0    0    0    0
```

0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0.2 0 0 0 0.8 0

0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0.1 0 0.3 0 0 0 0 0 0 0.7 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0.3 0 0 0 0 0 0 0.7
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0

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

	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0.7	0	0.1	0	0	0	0.2	0	0
	0	0	0	0	0	0	0					
0	0	0	0.6	0	0.6	0	0.3	0	0	0	0.1	0
	0	0	0	0	0	0	0					
0	0	0.7	0	0	0	0	0	0.3	0	0	0	
	0.1	0	0	0	0	0	0	0				
0	0.6	0	0	0	0	0	0	0.1	0.1	0	0	0
	0.2	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0.6	0	0	0	0	0	
	0.3	0	0	0	0.1	0	0	0				
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0.6	0	0	0	0.4	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0					
0	0	0	0	0	0	0	0	0	0	0	0	0

```
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0.7   0     0     0
      0.3   0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0.1   0     0     0     0     0
      0.7   0     0     0     0.2   0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     1     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     1     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     1     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0.2   0.6   0     0     0.2   0     0
0     0     0     0     0     0     0     0     0     0     0     0     0
      0     0     0     0     0     0     0

-1 -1 -1 -1 6 -1 -1 -1 -2 -1 -1 -4 -1 -1 -1 10 -1 -1 -1 -1
```

Output:

```
Value Function Matrix
-------------------------------------------------------------------
------
[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -
1.00000 -1.00000 -1.00000 ]

[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -
1.00000 -1.00000 -1.00000 ]
```

[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000]

[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000]

[6.00000 9.90000 12.63000 14.54100 16.14670 17.27133 18.06833 18.64976 19.05701 19.34323]

[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000]

[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000]

[-1.00000 -1.00000 -1.00000 -0.98000 -0.91200 -0.90560 -0.82624 -0.81435 -0.81232 -0.80652]

[-2.00000 -2.00000 -2.00000 -1.32000 -1.32000 -1.31040 -1.19280 -1.19261 -1.18958 -1.18135]

[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000]

[-1.00000 -1.00000 -1.00000 0.34000 0.34320 0.39200 0.50966 0.51086 0.51663 0.52489]

[-4.00000 -4.00000 -3.68000 0.24000 0.24640 0.34720 0.62173 0.62419 0.63674 0.65600]

[-1.00000 -1.00000 4.60000 4.60000 4.63200 5.02400 5.02464 5.03472 5.06217 5.06242]

[-1.00000 -1.00000 4.60000 4.60000 4.63200 5.02400 5.02464 5.03472 5.06217 5.06242]

[-1.00000 -1.00000 4.60000 4.60000 4.63200 5.02400 5.02464 5.03472 5.06217 5.06242]

[10.00000 10.00000 10.00000 10.00000 10.00000 10.00000 10.00000 10.00000 10.00000 10.00000]

[-1.00000 5.60000 5.60000 5.63200 6.02400 6.02464 6.03472 6.06217 6.06242 6.06367]

[-1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000 -1.00000]

```
[-1.00000 -1.00000 -1.00000 2.48000 2.48000 2.50560 2.81920 2.81971
2.82778 2.84974 ]

[-1.00000 -1.00000 -1.00000 1.36000 2.75200 2.77120 3.01664 3.14246
3.14872 3.16841 ]

Policy Matrix
------------------------------------------------------------------
------
[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 1 1 1 1 1 1 1 1 1 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 2 2 2 2 2 2 2 2 2 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 1 1 1 1 1 1 1 1 1 ]

[0 0 0 2 2 2 2 2 2 2 ]

[0 0 0 3 3 3 3 3 3 3 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 2 2 2 2 2 2 2 ]

[0 1 3 3 3 3 3 3 3 3 ]

[0 0 2 2 2 2 2 2 2 2 ]

[0 2 0 0 0 0 0 0 0 0 ]

[0 2 3 3 3 3 3 3 3 3 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 3 3 3 3 3 3 3 ]
```

```
[0 1 1 2 2 2 2 2 2 ]
```

Q3:

Input:
MDP1.txt

Output:

```
Value Function Matrix
-----------------------------------------------------------------------
------
[0.00000 1.00000 1.00000 1.03300 2.00000 2.00000 2.03300 3.00000
3.00000 3.03300 ]

[0.00000 0.00000 0.89440 0.99982 1.01836 1.89438 1.99982 2.01836
2.89438 2.99982 ]

[0.00000 0.00740 0.79868 0.89821 1.01255 1.79922 1.89819 2.01255
2.79922 2.89819 ]

[0.00000 0.00000 0.89440 0.90395 1.01836 1.89438 1.90393 2.01836
2.89438 2.90393 ]

[1.00000 1.00000 1.03300 2.00000 2.00000 2.03300 3.00000 3.00000
3.03300 4.00000 ]

[0.00000 0.00000 0.65950 0.89440 0.99982 1.65268 1.89438 1.99982
2.65243 2.89438 ]

[0.00000 0.65950 0.88056 0.98117 1.65268 1.87838 1.98044 2.65243
2.87829 2.98041 ]

[0.00000 0.00000 0.85218 0.89597 1.01254 1.85239 1.89576 2.01254
2.85239 2.89576 ]

[0.00000 0.03300 1.00000 1.00000 1.03300 2.00000 2.00000 2.03300
3.00000 3.00000 ]

[0.00000 0.89440 0.90395 1.01836 1.89438 1.90393 2.01836 2.89438
2.90393 3.01836 ]

Policy Matrix
-----------------------------------------------------------------------
------
```

```
[0 3 3 3 3 3 3 3 3 3 ]

[0 0 3 1 3 3 1 3 3 1 ]

[0 2 2 1 2 2 1 2 2 1 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 2 2 0 2 2 0 2 ]

[0 1 1 1 1 1 1 1 1 1 ]

[0 0 2 2 2 2 2 2 2 2 ]

[0 3 1 1 1 1 1 1 1 1 ]

[0 3 0 3 3 0 3 3 0 3 ]
```

Input:
MDP2.txt

Output:

```
Value Function Matrix
------------------------------------------------------------------
------
[0.00000 0.05690 1.00000 1.05685 2.00000 2.05687 3.00000 3.05687
4.00000 4.05687 ]

[0.00000 0.00000 0.99248 0.99496 1.99186 1.99518 2.99186 2.99522
3.99186 3.99523 ]

[0.49310 1.25265 1.59912 2.29697 2.61933 3.30618 3.62356 4.30811
4.62444 5.30851 ]

[0.00000 0.00060 0.06647 0.99360 1.04189 1.99294 2.04200 2.99284
3.04193 3.99274 ]

[0.00000 0.57150 0.99995 1.57148 1.99997 2.57149 2.99997 3.57149
3.99997 4.57149 ]

[0.00000 1.00000 1.00023 2.00000 2.00027 3.00000 3.00028 4.00000
4.00028 5.00000 ]
```

```
[1.00000 1.00000 2.00000 2.00023 3.00000 3.00027 4.00000 4.00028
5.00000 5.00028 ]

[0.00000 0.07750 0.08162 0.96575 1.08132 1.96484 2.08131 2.96485
3.08131 3.96485 ]

[0.00000 0.01119 0.57150 1.00000 1.57148 2.00000 2.57149 3.00000
3.57149 4.00000 ]

[0.00000 0.99995 1.00003 1.99996 2.00026 2.99996 3.00030 3.99996
4.00031 4.99996 ]

Policy Matrix
------------------------------------------------------------------
------
[0 3 0 3 0 3 0 3 0 3 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 0 0 0 0 0 0 0 0 0 ]

[0 1 2 1 1 1 1 1 1 1 ]

[0 2 1 2 1 2 1 2 1 2 ]

[0 2 0 2 0 2 0 2 0 2 ]

[0 0 2 2 2 2 2 2 2 2 ]

[0 3 1 1 1 1 1 1 1 1 ]

[0 3 1 0 1 0 1 0 1 0 ]

[0 2 2 2 2 2 2 2 2 2 ]
```