PROBLEM 3
PART 1

MARKOV DECISION PROBLEM (MDP)
VALUE ITERATIONS

Problem 1: Value Iteration for both version of Magneto on MDP

[Please use Google Colab or Jupyter notebook to run submitted ipython notebook.]

Maximum expectation possible: +20
Minimum expectation possible: -20

On the basis of the random moves of both jean and magneto(case 1), the expectation table will be created. [Each output test case is run for 200 iterations if it does not converge before 200 iterations]

Note: Coordinates of the board are considered as given below (x_cordinate,y_cordinate):

```
(0, 0)  |(0, 1)  |(0, 2)  |(0, 3)  |(0, 4)  |
(1, 0)  |(1, 1)  |(1, 2)  |(1, 3)  |(1, 4)  |
(2, 0)  |(2, 1)  |(2, 2)  |(2, 3)  |(2, 4)  |
(3, 0)  |(3, 1)  |(3, 2)  |(3, 3)  |(3, 4)  |
(4, 0)  |(4, 1)  |(4, 2)  |(4, 3)  |(4, 4)  |
```

U,D,L,R,X stands for move up,down,left,right and no move in the board output.

Output:

```
Initial rewards configurations:
-----------------------------------
 0.00| 0.00| 0.00| 0.00| 0.00|
-----------------------------------
 0.00| 0.00| 0.00| 0.00| 0.00|
-----------------------------------
 0.00| 0.00| 0.00| 0.00| 0.00|
-----------------------------------
 0.00|-20.00| 0.00| 0.00| 0.00|
-----------------------------------
 0.00| 20.00| 0.00| 0.00| 0.00|


Initial random policy
------------------------------
```

```
 D  |  D  |  R  |  R  |  D  |
-------------------------------
 L  |  U  |  D  |  U  |  U  |
-------------------------------
 R  |  L  |  X  |  R  |  L  |
-------------------------------
 D  |  R  |     |  X  |  L  |
-------------------------------
 R  |  U  |  U  |  X  |  L  |
```

Initial random probability

```
-----------------------------------
 0.81| 0.85| 0.04| 0.19| 0.22|
-----------------------------------
 0.91| 0.00| 0.66| 0.69| 0.50|
-----------------------------------
 0.79| 0.61| 0.82| 0.32| 0.80|
-----------------------------------
 0.38| 0.73| 0.00| 0.08| 0.03|
-----------------------------------
 0.87| 0.36| 0.99| 0.22| 0.30|
```

After 200 value iterations:

Final Values:

```
-----------------------------------
 12.28| 14.45| 17.00| 17.00| 20.00|
-----------------------------------
 10.44| 12.28| 14.45| 17.00| 17.00|
-----------------------------------
 10.44| 10.44| 12.28| 14.45| 17.00|
-----------------------------------
 17.00| 14.45| 0.00| 14.45| 14.45|
-----------------------------------
 20.00| 17.00| 17.00| 17.00| 14.45|
```

Final Policy (D,U,L,R,X for down,up,left,right and no move:)

```
-------------------------------
 D  |  R  |  R  |  R  |  X  |
-------------------------------
 U  |  U  |  U  |  U  |  U  |
-------------------------------
 D  |  U  |  U  |  U  |  U  |
-------------------------------
  D  |  D  |     |  D  |  U  |
-------------------------------
  X  |  L  |  L  |  L  |  L  |
```

FOR ACTIVE MAGNETO MOVE OUTPUTS:

Initial rewards configuration:

```
-----------------------------------
 0.00| 0.00| 0.00| 0.00| 0.00|
-----------------------------------
 0.00| 0.00|-20.00| 0.00| 0.00|
-----------------------------------
 0.00| 0.00| 0.00| 0.00| 0.00|
-----------------------------------
 0.00| 0.00| 0.00| 0.00| 0.00|
-----------------------------------
 0.00| 20.00| 0.00| 0.00| 0.00|
```

Initial random policy

```
-------------------------------
  X  |  U  |  R  |  R  |  U  |
-------------------------------
  L  |  D  |  X  |  U  |  R  |
-------------------------------
  X  |  D  |  X  |  R  |  U  |
-------------------------------
  D  |  R  |     |  X  |  L  |
-------------------------------
  X  |  L  |  D  |  D  |  R  |
```

Initial random probability

```
-----------------------------------
 0.52| 0.63| 0.39| 0.61| 0.80|
-----------------------------------
 0.31| 0.43| 0.81| 0.77| 0.55|
-----------------------------------
 0.42| 0.97| 0.83| 0.97| 0.78|
-----------------------------------
```

```
 0.92| 0.25| 0.00| 0.15| 0.80|
-----------------------------------
 0.39| 0.73| 1.00| 0.40| 0.31|
```

After 200 value iterations:

```
Final Values:
-----------------------------------
 10.44| 12.28| 14.45| 20.00| 20.00|
-----------------------------------
 10.44| 14.45| 12.28| 14.45|-17.00|
-----------------------------------
 14.45| 17.00| 12.28| 12.28| 3.94|
-----------------------------------
 17.00| 20.00| 0.00|-20.00| 0.00|
-----------------------------------
 0.00| 17.00| 20.00| 14.45| 0.00|

Final Policy (D,U,L,R,X  for down,up,left,right and no move:)
-------------------------------
  R  |  R  |  L  |  R  |  X  |
-------------------------------
  D  |  D  |  U  |  U  |  U  |
-------------------------------
  D  |  D  |  L  |  U  |  X  |
-------------------------------
  R  |  D  |     |  U  |  U  |
-------------------------------
  U  |  X  |  L  |  L  |  L  |
```