Artificial Intelligence Project 3: MDPs

Our Problem: We are given a matrix which contains final values, blank spaces which need to be calculated and walls. Our target is to calculate all the best value(best direction) in each blank space.

Algorithms: I calculate the final value use MDP algorithm, which can be divided to two different methods : value iteration and policy iteration. Although the basic theory between these two are same, the way to calculate is different. Value iteration calculate all the directions values each time until converge, the policy iteration only calculate one direction value each time, after converge we choose the best direction and then do it again and again.

I have submitted the source code, I wrote comment in every place which if it is necessary.

Result comparison : Both of they can get the optimal result, So I use these two results to check whether I wrote the right code. I recorded iterate times until we get the optimal result.

Time complexity comparison: Value iteration take O(S * S * A) (S means the matrix's length, A means the sum of directions each position can go) time in the worst case and policy iteration takes O(S * S * T)(T means the total times we change the direction). It means that when A is vary large so the policy iteration is A/T times faster than value iteration.

This is my running record.

```
Graph 1 :
ValueIteration Result :
24.18 27.46 28.41 1.00 1.98 1.99 2.39
24.82 31.57 35.90 -1.00 2.38 1.00 2.77
-1.00 36.08 46.04 -1.00 3.31 4.00 3.38
66.55 1.00 59.57 -1.00 2.38 1.00 2.77
83.99 100.00 82.73 -100.00 2.60 3.00 2.64
66.65 2.00 59.66 -1.00 2.56 3.00 2.59
0.00 36.73 46.17 -1.00 1.84 1.00 2.15
Times Running:25 * 4

PolicyIteration Result :
24.18 27.46 28.41 1.00 1.98 1.99 2.39
24.82 31.57 35.90 -1.00 2.38 1.00 2.77
-1.00 36.08 46.04 -1.00 3.31 4.00 3.38
66.55 1.00 59.57 -1.00 2.38 1.00 2.77
83.99 100.00 82.73 -100.00 2.60 3.00 2.64
66.65 2.00 59.66 -1.00 2.56 3.00 2.59
0.00 36.73 46.17 -1.00 1.84 1.00 2.15
Times Running:119
```

```
Graph 2 :
ValueIteration Result :
0.15 0.29 0.52 1.00 0.69 0.67
0.01 0.15 0.08 -1.00 0.76 1.00
-1.00 0.27 0.12 -1.00 1.77 4.00
0.56 1.00 0.48 -1.00 0.82 1.00
0.97 2.00 0.88 -1.00 1.35 3.00
0.00 0.96 0.53 -1.00 0.64 1.00
Times Running:19 * 4

PolicyIteration Result :
0.15 0.29 0.52 1.00 0.69 0.67
0.01 0.15 0.08 -1.00 0.76 1.00
-1.00 0.27 0.12 -1.00 1.77 4.00
0.56 1.00 0.48 -1.00 0.82 1.00
0.97 2.00 0.88 -1.00 1.35 3.00
0.00 0.96 0.53 -1.00 0.64 1.00
Times Running:45
```

Performance Measure: In small matrix or small directions we can choose, sometimes Value iteration faster than Policy iteration. But when the matrix become more and more complex or have many directions can be chose. policy iteration is faster than value iteration. And I believe that in complex situation, policy iteration is better choice.