

حل مسئله‌ی راهروی مارپیچ با برنامه‌سازی پویا

امید صادق‌نژاد
۹۹۱۲۶۵۴

نتایج کد پیاده‌سازی شده برای مسئله‌ی راهروی مارپیچ به صورت زیر می‌باشد. حرکت‌ها به این صورت می‌باشند
{0:up, 1:down, 2:right, 3:left}

برای حالتی که هر خانه امتیاز -0.04 باشد: مقادیر سیاست‌ها و ارزش هر خانه در ابتدای الگوریتم به صورت زیر است.

```
1 [ 1.  2.  2.  1.]
2 [ 1. nan 0.  2.]
3 [ 0.  0.  2.  1.]]
4
5 [[0.63945675 0.71107754 0.62066005 0.
6  [0.87689969          nan 0.66123633 0.
7  [0.51633973 0.53541164 0.41536241 0.51248974]]
8
```

این مقادیر در انتهای آموزش الگوریتم به صورت زیر می‌باشد.

```
1 [[0.92 0.96 1.  0. ]
2  [0.88 nan 0.96 0. ]
3  [0.84 0.88 0.92 0.88]]
4
5 [[ 2.  2.  2.  1.]
6  [ 0. nan 0.  2.]
7  [ 2.  2.  0.  3.]]
8
```

برای حالتی که هر خانه امتیاز -4 باشد:

```
1 [[ 1.  0.  2.  2.]
2  [ 1. nan 0.  1.]
3  [ 1.  1.  2.  1.]]
4
5 [[0.45017349 0.49564106 0.84728801 0.
6  [0.39408231          nan 0.16206914 0.
7  [0.60292376 0.46087325 0.50513331 0.53498239]]
8
```

این مقادیر در انتهای آموزش الگوریتم به صورت زیر می‌باشد.

```
1 [[ -7.  -3.  1.  0.]
2  [-11. nan -1.  0.]
3  [-13. -9. -5. -1.]]
4
5 [[ 2.  2.  2.  2.]
6  [ 0. nan 2.  1.]
7  [ 2.  2.  2.  0.]]
8
```

در حالت دوم عامل سعی دارد سریع‌تر به پایان برسد.