

动态规划算法实验报告

Arthals 2110306206

一、实验要求

(1) 阅读《基于动态规划的路径规划-实验指导书》，尝试运行并理解动态规划算法在冰湖路径规划问题上的示例代码。

(2) 在理解示例代码的基础上补全代码中的空缺，运行实验并记录算法收敛所需要的迭代次数。

二、代码填空内容

1. Policy Evaluation

```
# TODO: 计算用于更新  $V[s]$  的  $v_{\text{new}}$ 
v_new += policy[s][a] * prob * (reward + gamma * V[next_state])
```

2. Policy Iteration

```
# TODO: 计算动作  $a$  的价值  $A[a]$ 
A[a] += prob * (reward + gamma * V[next_state])
```

3. Value Iteration

```
# TODO: 计算动作  $a$  的价值  $A[a]$ 
A[a] += prob * (reward + gamma * V[next_state])
```

三、实验结果展示

```
arthals in 🍷 Aurora in Fundamentals-Of-Artificial-Intelligence-2024Spring-PKU/Homework/08-Reinforcement-Learning via 🐍 v3.11.8 via 🍷 torch ⚡ took 6s759ms
at 14:15:34 > python dynamic_programming.py
2
57
策略迭代算法与价值迭代算法的最终策略一致。
```