|  |  |
| --- | --- |
| **自然语言处理及应用**  **实验报告** | |
|  | |
|  |  |
| **名称** | 最小编辑距离 |
| **姓名** | 郭淇 |
| **班级** | 软件72班 |
| **学号** | 2174111345 |
| Email | 1325930068@qq.com |
| **日期** | 2021-3-10 |

# 实验目的

实现英文和中文字符串的最小编辑距离，最小编辑路径的计算并显示结果。可以输入不同字符串；可以通过修改操作代价来改变最小编辑距离。

# 实验环境

Python3.8

Windows10

Pycharm编辑器

# 实验方法

编辑操作包括增加、删除和替换，采用动态规划的思想，可以得到以下的表格：

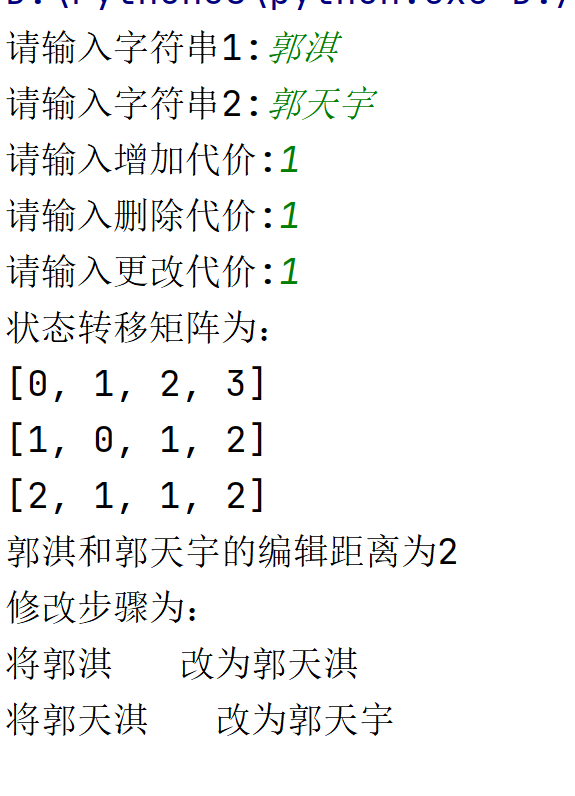
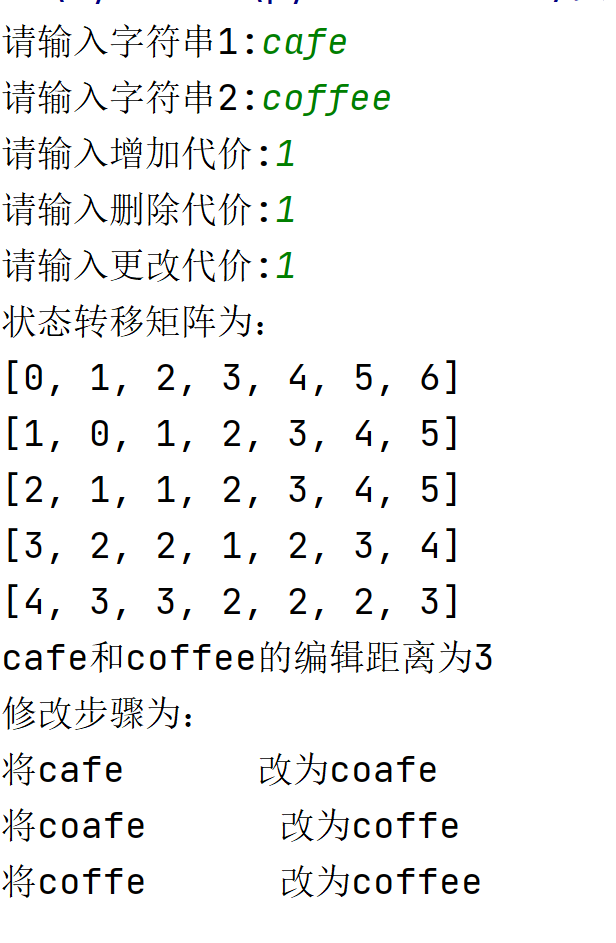
|  |  |  |
| --- | --- | --- |
| **矩阵方向** | **编辑操作** | **计算公式(D(i,j)更新为)** |
| 向右走 | 删除 |  |
| 向下走 | 插入 |  |
| 对角线 | 替换/匹配 |  |

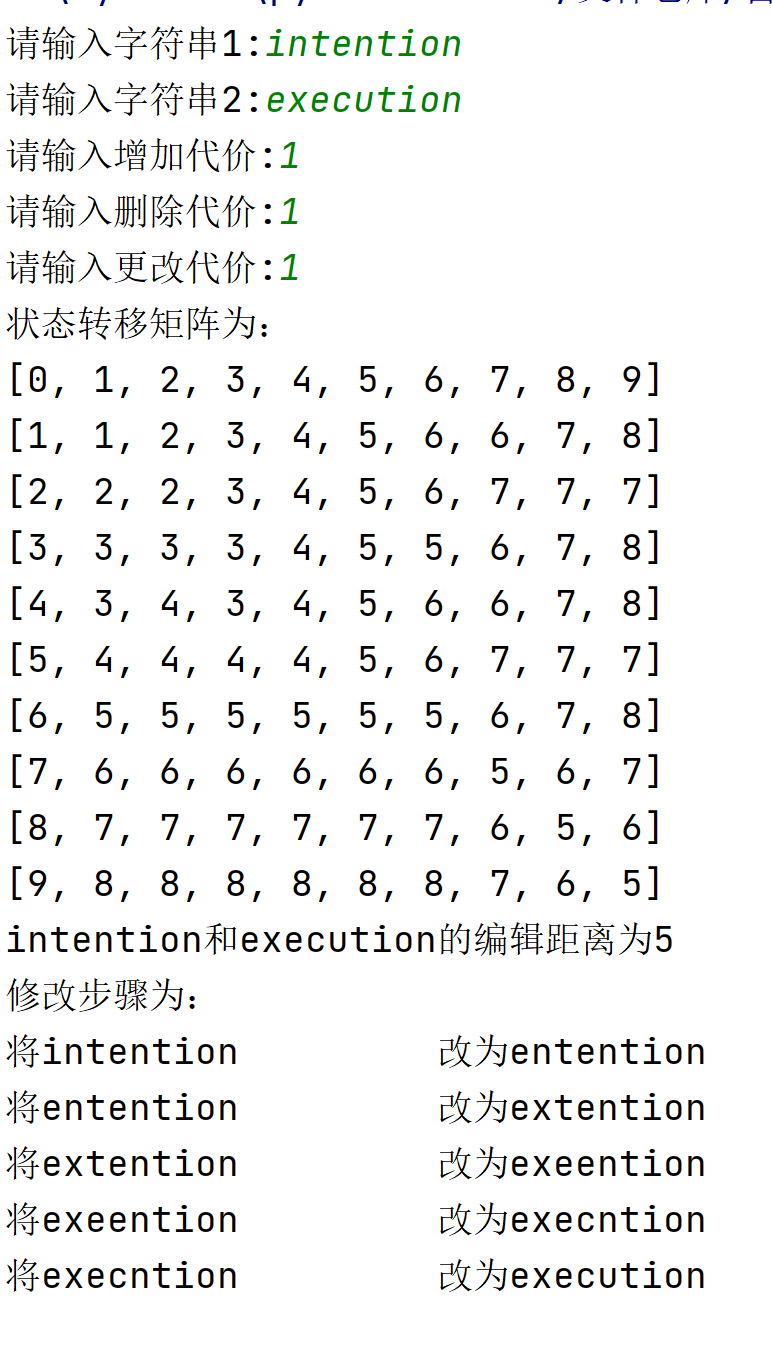
更新公式为：

回溯时记住不同操作的顺序，根据不同的操作更改字符串。可能会遇见多种情况都成立的情况存在，我只考虑了其中一种情况。

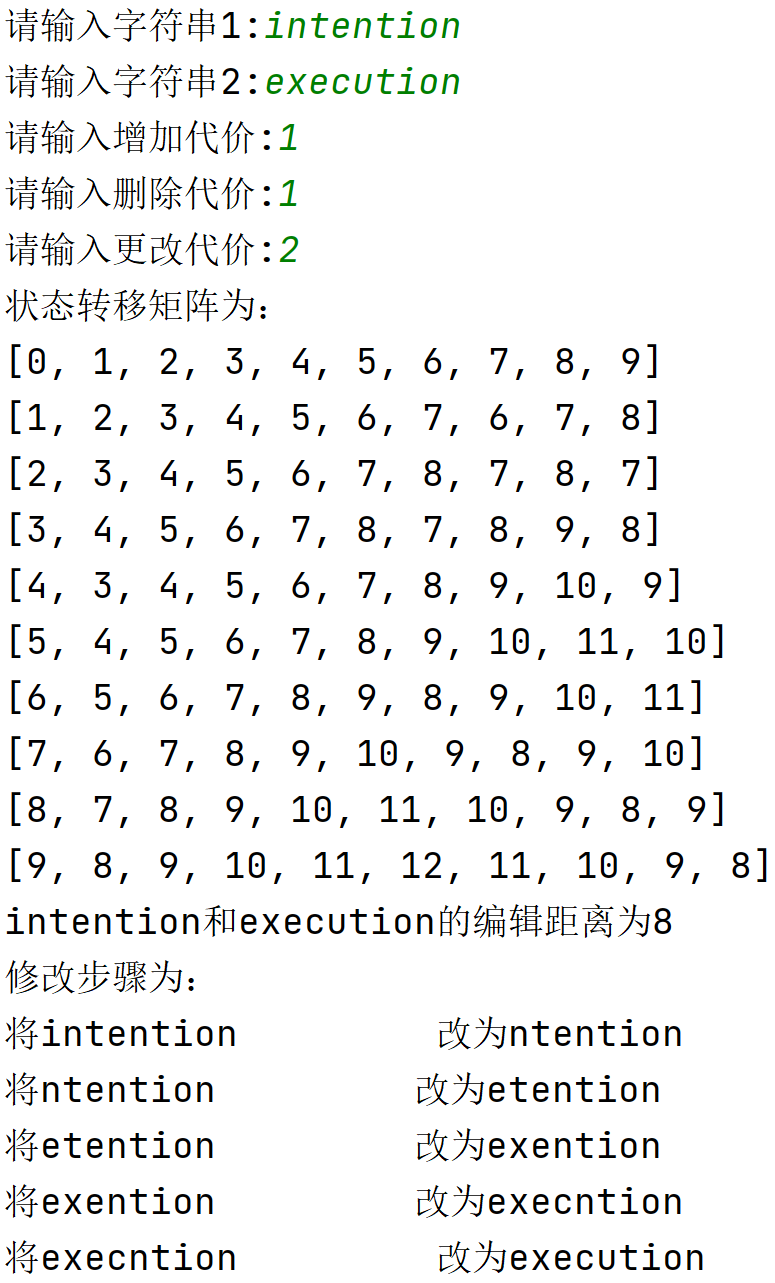
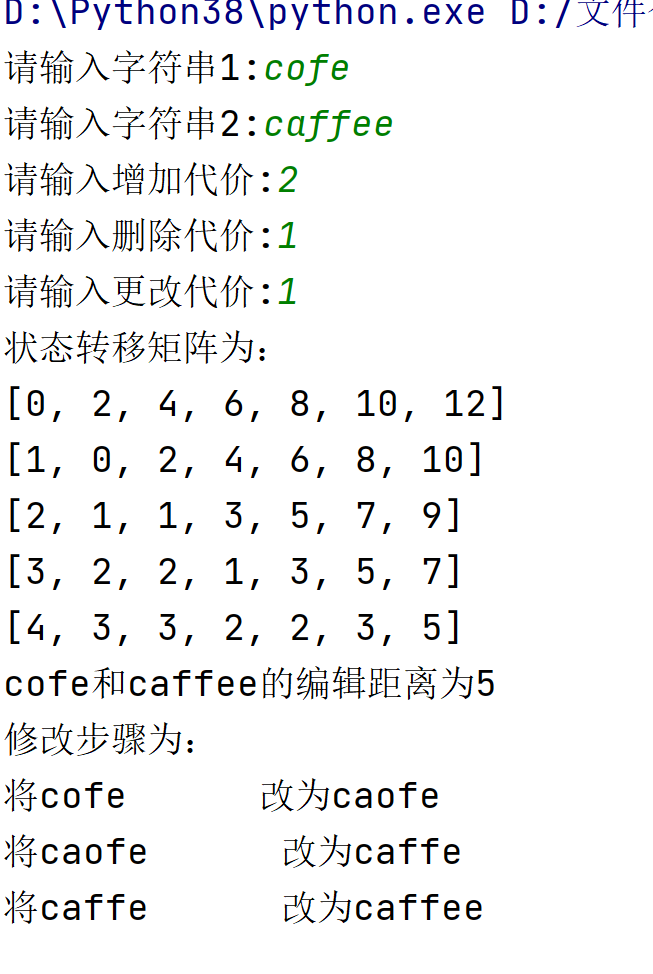
1. **实验结果**

代价都为1时：

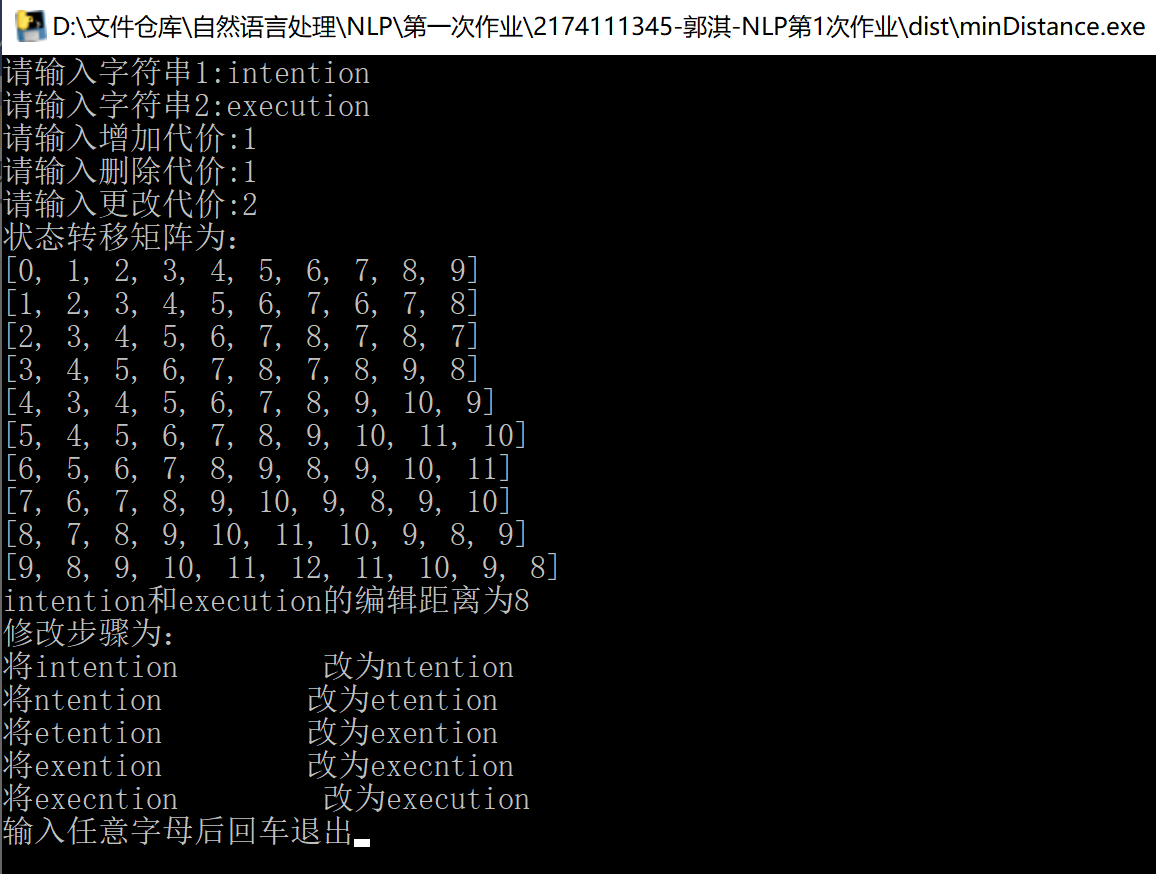




更改增加、删除或修改操作的代价



可执行文件执行效果：



1. **遇到问题及解决思路**

在回溯矩阵的时候可能会遇到增加、删除或替换代价相同的情况。

我采用的方法是只考虑其中的一种情况，可能会有更好的方法将所有的情况输出，但我没有实现。