|  |  |
| --- | --- |
| **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** | 2021-3-7 |

# 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

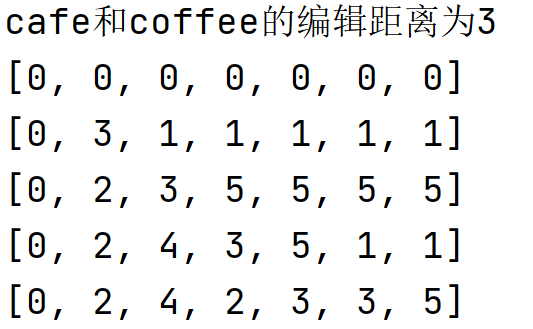
编辑操作包括增加、删除和替换，采用动态规划的思想，可以得到以下的表格：

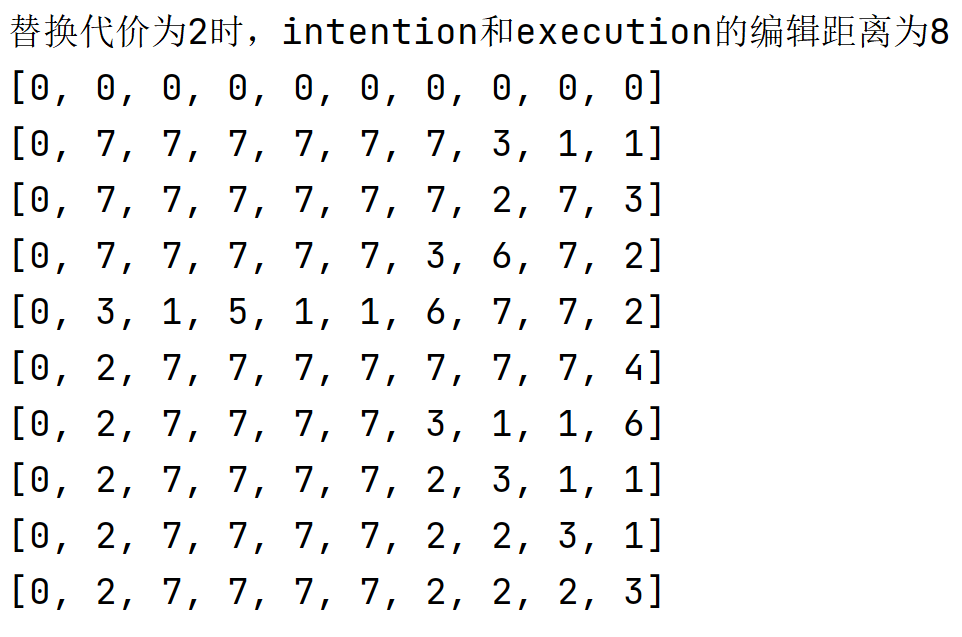
|  |  |  |
| --- | --- | --- |
| **矩阵方向** | **编辑操作** | **计算公式(D(i,j)更新为)** |
| 向右走 | 删除 |  |
| 向下走 | 插入 |  |
| 对角线 | 替换/匹配 |  |

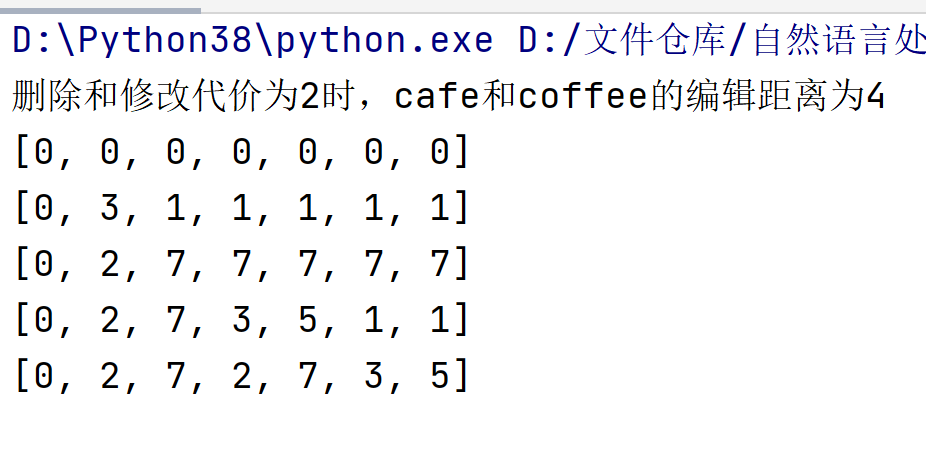
更新公式为：

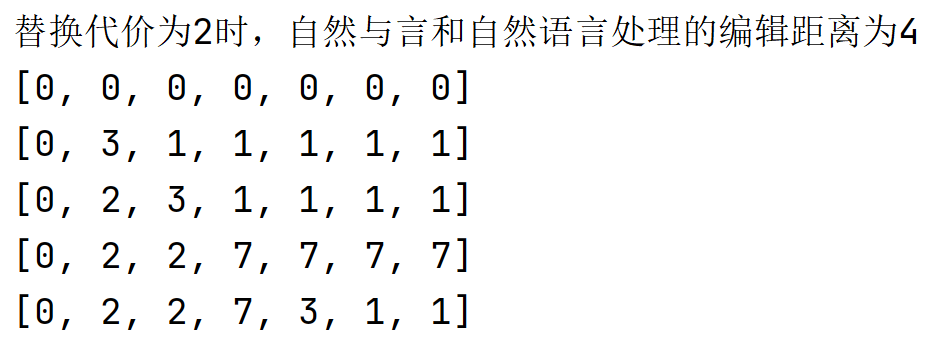
1. **实验结果**

下面的矩阵表示回溯矩阵。1表示经过增加操作，2表示经过删除操作，3表示经过替换操作，4表示可能是删除或替换操作，5表示可能是增加或替换操作，6表示可能是增加或删除操作，7表示三种操作都有可能。







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。我采用的是将7种可能的情况都表示出来，但这种方法不太好，不过没有想到更好的方法。