# 最短编辑距离

## 1　题目描述

UNIX系统下有一个行编辑器ed，它每次只对一行文本做删除一个字符、插入一个字符或替换一个字符三种操作。例如某一行的内容是“ABC”，经过把第二个字符替换成“D”、删除第一个字符、末尾插入一个字符“B”，这三步操作后，内容就变成了“DCB”。即“ABC”变成“DCB”需要经过3步操作，我们称它们的编辑距离为3。

现在给你两个任意字符串（不包含空格），请帮忙计算它们的最短编辑距离。

### 1.1　输入描述:

输入包含多组数据。每组数据包含两个字符串m和n，它们仅包含字母，并且长度不超过1024。

### 1.2　输出描述:

对应每组输入，输出最短编辑距离。

### 1.3　输入例子:

ABC CBCD

ABC DCB

### 1.4　输出例子:

2

3

## 2　解题思路

设A和B是2个字符串。要用最少的字符操作将字符串A转换为字符串B。这里所说的字符操作包括:

1. 删除一个字符;
2. 插入一个字符；
3. 将一个字符改为另一个字符。

将字符串A变换为字符串B所用的最少字符操作数称为字符串A到B的编辑距离。设A的长度为m，B的长度为n创建一个二维数组d，大小为(m+1)\*(n+1)，来记录a1-am与b1-bn之间的编辑距离，要递推时，需要考虑对其中一个字符串的删除操作、插入操作和替换操作分别花费的开销，从中找出一个最小的开销即为所求结果。

操作步骤：

1. 情况一：当A的长度为0，B的长度为j时，最小编辑距离就是j。
2. 情况二：当A的长度为i，B的长度为0时，最小编辑距离就是i。
3. 情况三：当A的长度为i，B的长度为j时，d[i][j]=min{d[i-1][j]+1,d[i][j-1]+1, d[i-1][j-1]+(A[i]==B[j]?0:1)}

其中：

* d[i][j]表示A的前i个字符和B的前j个字符相同后的最短距离。
* dp[i][j]来自于三种状态
* 删除，d[i-1][j]+1，A1,…,Ai-1经过操作可以变成了B1,…,Bj，那么A1,…,Ai-1Ai变为B1,…,Bj一定要删除Ai。
* 插入，d[i][j-1]+1，A1,…,Ai经过操作可以变成了B1,…,Bj-1Bj，那么A1,…,Ai-1Ai变为B1,…,Bj-1Bj一定要添加一个字符。
* 替换，如果A[i]=B[j]，可以不进行额外的操作，那么有d[i][j]=d[i-1][j-1]，如果不A[i]≠B[j]，那么就要进行一次替换操作，有d[i][j]=d[i-1][j-1]+1。

**注意**：此处字符串中字符开始的下标从1开始计算

根据分析可以得到递推方程：