

算法实现题 3-13 正则表达式匹配问题（习题 3-22）

★问题描述：

许多操作系统采用正则表达式实现文件匹配功能。一种简单的正则表达式由英文字母、数字及通配符“*”和“？”组成。“？”代表任意一个字符。“*”则可以代表任意多个字符。现要用正则表达式对部分文件进行操作。

试设计一个算法，找出一个正则表达式，使其能匹配的待操作文件最多，但不能匹配任何不进行操作的文件。所找出的正则表达式的长度还应是最短的。

★编程任务：

对于给定的待操作文件，找出一个能匹配最多待操作文件的正则表达式。

★数据输入：

由文件 input.txt 提供输入数据。文件由 n ($1 \leq n \leq 250$) 行组成。每行给出一个文件名。文件名由英文字母和数字组成。英文字符要区分大小写，文件名长度不超过 8 个字符。文件名后是一个空格符和一个字符“+”或“-”。“+”表示要对该行给出的文件进行操作，“-”表示不进行操作。

★结果输出：

程序运行结束时，将计算出的最多文件匹配数和最优正则表达式输出到文件 output.txt 中。文件的第 1 行中的数是计算出的最多文件匹配数。文件的第 1 行是最优正则表达式。

输入文件示例

```
input.txt
EXCHANGE +
EXTRA +
HARDWARE +
MOUSE -
NETWORK -
```

输出文件示例

```
output.txt
3
*A*
```