## 实验报告(四)

实验内容

**利用模式匹配算法，实现课件中的案例 4.1： 病毒感染检测。**

**问题描述如下：**

研究者将人的 DNA 和病毒 DNA 均表示成由一些字母组成的字符串序列。

然后检测某种病毒 DNA 序列是否在患者的 DNA 序列中出现过，如果出现过，则此人感染 了该病毒，否则没有感染。

例如，假设病毒的 DNA 序列为 baa，患者 1 的 DNA 序列为 aaabbba，则感染，患者 2 的 DNA 序列为 babbba，则未感染。（注意，人的 DNA 序列是线性的，而病毒的 DNA 序列是环状的)。

要求：相关信息从文件"病毒感染检测输入数据.txt"读入，将检测结果写入到文件"病 毒感染检测输出结果.txt"中。

**提高内容： 利用 KMP 算法进行匹配。算法流程**

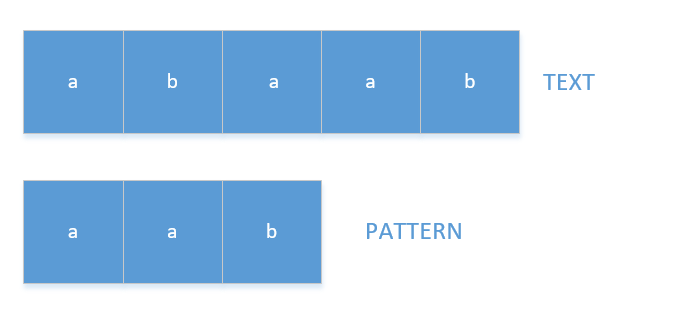
算法流程

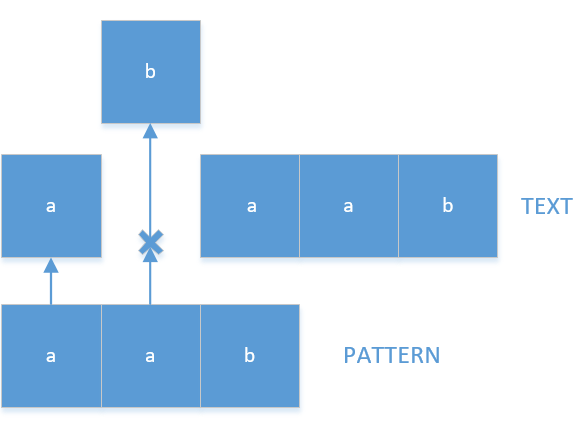
**循环模式串的处理方式**

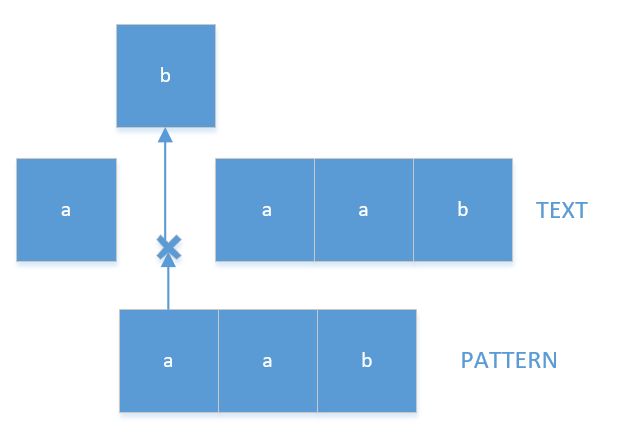
若原始模式串长度为m，将模式串复制一倍，从以第一个元素为头，到以下一个元素为头遍历m次，每次遍历中达到长度m即为完成匹配，最后即遍历了循环模式串的所有结果

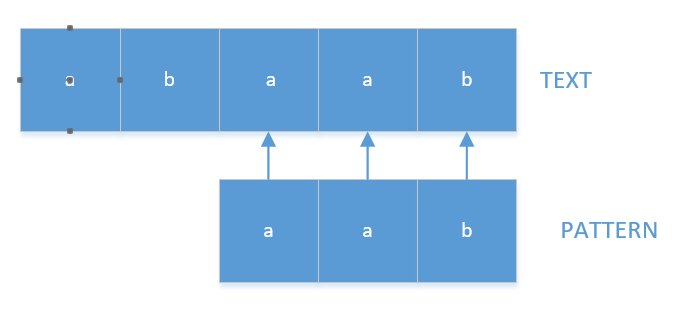
**BF算法（以模式串aab，文本串abaab）**

依次匹配，若不匹配，模式串从文本串下一个位置重头开始匹配



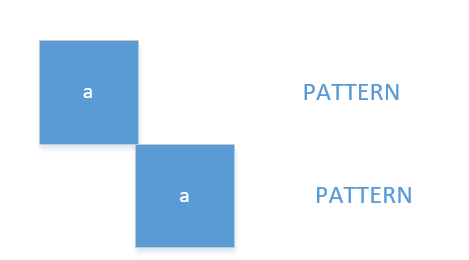


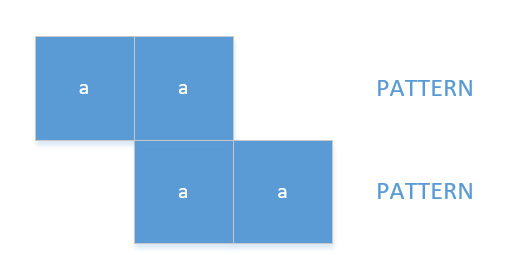


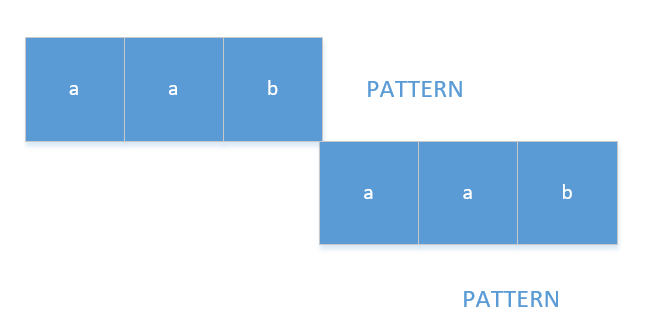


**KMP算法（以模式串aab，文本串abaab）**

1.求模式串的next数组，即找最大部分，下一个匹配错误时退回的最好位置

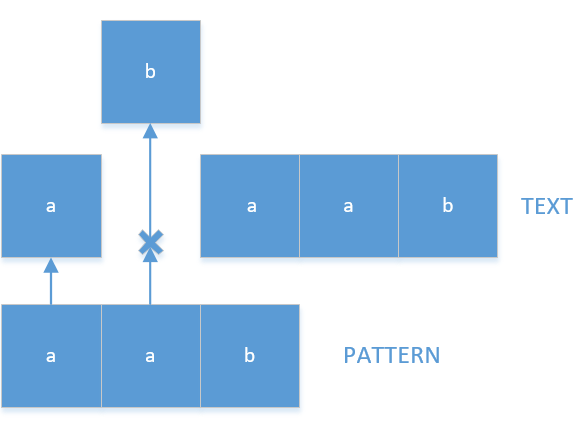


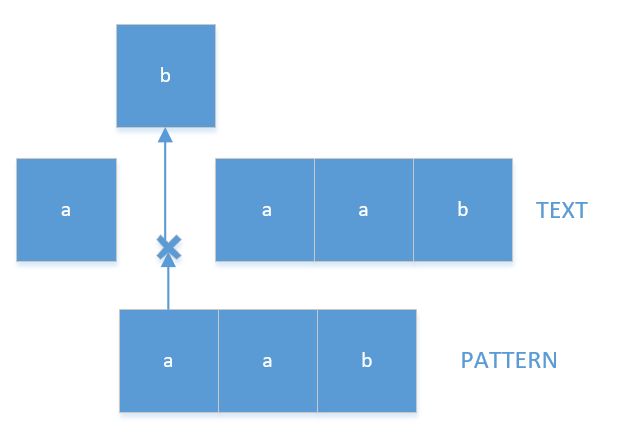


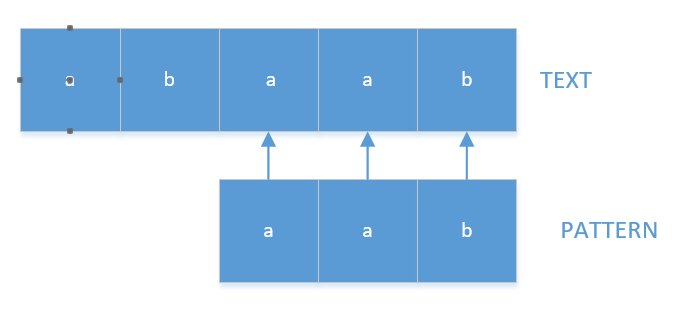


next数组值分别为0，1，0

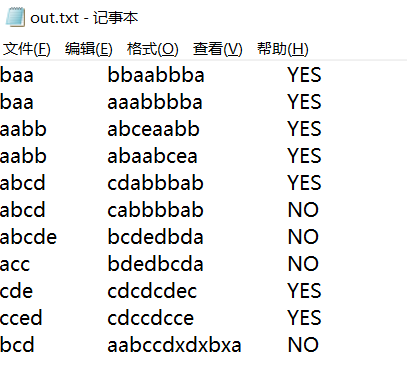
2.KMP







实验结果



实验小结

1. 学会了BF字符串模式匹配算法的写法
2. 学会了KMP字符串模式匹配算法的写法
3. 学会了若模式串是循环串的处理方法
4. 学会了用重定向的方式输入输出文件数据