

第四章 字符串

1. 求下列字符串的 next 数组。

$t = \text{'abcdaabcddaba'}$

$s = \text{'XYYXYZXZXYYXZ'}$

2. 请设计算法消除字符串中的 b 和 ac, 需要满足条件:

(1)字符串只能遍历一次。

(2)不能够使用额外空间, 即 $O(1)$ 的空间。

如 “acbac” \rightarrow “ ” , “ababac” \rightarrow “aa” .

2. 求已知长度为 m 的字符串 P 和长度为 n 的字符串 T , $n > m$, $S = PT$ (由 P 和 T 连接形成的长度为 $m+n$ 的字符串), 类似于 next 数组, 在 S 上定义 nextset 数组, 定义 $\text{nextset}[i]$ 如下:

$$\text{nextset}[i] = \begin{cases} \{-1\}, & \text{if } i = 0 \\ \{k | 0 < k < i, S[0 \dots (k-1)] = S[(i-k) \dots (i-1)], \text{ if } k \text{ exists}\} \\ \{0\}, & \text{otherwise} \end{cases}$$

分两种情况, 说明如何通过检查 nextset 数组来判断 P 是否在文本 T 中出现。

1) 若 P 不会出现在 T 尾部, 即 $P \neq T[(n-m-2) \dots (n-1)]$, 写出充要条件并证明。

2) 若 P 可能出现在 T 尾部, 写出充要条件但不需证明。

3. 求给定一个长度为 n 的字符串 S 和一个整数 i , 要求设计一个线性时间, 常数空间的算法, 将 $S[0...i]$ 移动到 $S[(n-i-1)...(n-1)]$, 将 $S[(i+1)...(n-1)]$ 移动到 $S[0...(n-i-2)]$ 。如输入 “HELLOWORLD”, $i=4$, 输出 “WORLDHELLO”。