

字符串算法选讲

l0nl1f3

福州第三中学

2017年6月

Tips

- ▶ 忘了Tips要写什么了

KMP

- ▶ 给定一个长度为 n 字符串 S ,给定一个长度为 m 字符串 T ,问 T 在 S 中出现了几次

KMP

- ▶ 给定一个长度为 n 字符串 S ,给定一个长度为 m 字符串 T ,问 T 在 S 中出现了几次
- ▶ $n \leq m \leq 5 * 10^6$

KMP

- ▶ 朴素匹配,枚举起点 l ,比较 $s_{l...l+m-1}$ 和 T 是否相等,遇到不等(失配)就移动 l
- ▶ 复杂度 $O(nm)$

KMP

- ▶ 朴素匹配,枚举起点 l ,比较 $s_{l...l+m-1}$ 和 T 是否相等,遇到不等(失配)就移动 l
- ▶ 复杂度 $O(nm)$
- ▶ KMP(Knuth–Morris–Pratt algorithm),俗称看毛片算法

KMP

- ▶ 朴素匹配,枚举起点 l ,比较 $s_{l\dots l+m-1}$ 和 T 是否相等,遇到不等(失配)就移动 l
- ▶ 复杂度 $O(nm)$
- ▶ KMP(Knuth–Morris–Pratt algorithm),俗称看毛片算法
- ▶ 定义一个串的border为满足 $s_{1\dots x} = s_{n-x+1\dots n}$ 的前缀

KMP

- ▶ 朴素匹配,枚举起点 l ,比较 $s_{l...l+m-1}$ 和 T 是否相等,遇到不等(失配)就移动 l
- ▶ 复杂度 $O(nm)$
- ▶ KMP(Knuth–Morris–Pratt algorithm),俗称看毛片算法
- ▶ 定义一个串的border为满足 $s_{1...x} = s_{n-x+1...n}$ 的前缀
- ▶ KMP的next数组存储的是 T 串每个前缀的最大border的长度

KMP

- ▶ 失配时按border跳

KMP

- ▶ 失配时按border跳
- ▶ 构建next数组 $O(n)$ ，匹配总复杂度 $O(n)$