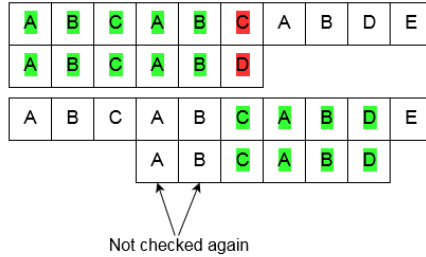


讨论 / 学习分享 / 分享 | 【题单】字符串 (KMP...

## 分享 | 【题单】字符串 (KMP/Z函数/Manacher/字符串哈希/AC自动机/后缀数组/子序列自动机)

+ 关注 TA

灵茶山艾府 发起于 2024-10-07 · 最近编辑于 9 天前 · 来自浙江



## 一、KMP

定义  $s$  的真前缀为不等于  $s$  的前缀,  $s$  的真后缀为不等于  $s$  的后缀。

定义  $s$  的 border 为既是  $s$  的真前缀又是  $s$  的真后缀的字符串。例如在  $s = aabcaa$  中,  $a$  和  $aa$  都是  $s$  的 border。

对于模式串  $p$  的每个前缀  $p[:i]$ , 计算这个前缀的最长 border 长度, 记在  $\pi$  数组中。

利用  $\pi$  数组, 可以快速计算模式串  $p$  出现在文本串  $t$  的哪些位置上。代码模板见 [我的题解](#)。

注:  $\pi$  数组的定义参考《算法导论》, 国内数据结构教材通常定义为  $next$  数组。以严蔚敏那本为例, 二者的关系是  $next[i+1] = \pi[i] + 1$ , 即  $\pi$  数组整体右移一位, 元素值加一。

- 28. 找出字符串中第一个匹配项的下标 模板题
- 796. 旋转字符串 做到  $O(n+m)$
- 1392. 最长快乐前缀 1876
- 3036. 匹配模式数组的子数组数目 II 1895
- 1764. 通过连接另一个数组的子数组得到一个数组 做到线性时间复杂度
- 1668. 最大重复子字符串 做到  $O(n+m)$
- 459. 重复的子字符串 做到  $O(n)$
- 3008. 找出数组中的美丽下标 II 2016
- 214. 最短回文串 也可以用 Manacher 算法
- 686. 重复叠加字符串匹配 ~2200
- 1397. 找到所有好字符串 2667
- 3037. 在无限流中寻找模式 II (会员题) 同 28 题

## 二、Z 函数 (扩展 KMP)

对于字符串  $s$ , 定义  $z[i]$  表示后缀  $s[i:]$  与  $s$  的 LCP (最长公共前缀) 的长度, 其中  $s[i:]$  表示从  $s[i]$  到  $s[n-1]$  的子串。

常用技巧是构造字符串  $pattern + s$ , 如果发现  $z[m+i] \geq m$  ( $m$  是  $pattern$  的长度), 则说明从  $s[i]$  开始的子串与  $pattern$  匹配。

所以上面的一些 KMP 题目 (子串匹配相关的), 也可以用 Z 函数解决。读者可以试试用 Z 函数解决 28. 找出字符串中第一个匹配项的下标。

- 2223. 构造字符串的总得分 2220 模板题
- 3031. 将单词恢复初始状态所需的最短时间 II 2278
- 3045. 统计前后缀下标对 II 2328
- 3303. 第一个几乎相等子字符串的下标 2509
- 3292. 形成目标字符串需要的最少字符串数 II 2662

## LCP 数组

- 2430. 对字母串可执行的最大删除数 2102
- 3388. 统计数组中的美丽分割 ~2350
- 1977. 划分数字的方案数 2817

## 三、Manacher 算法

Manacher 算法可以计算以  $s[i]$  (或者  $s[i]$  和  $s[i+1]$ ) 为回文中心的最长回文子串的长度。

此外, 还可以:

- 判断任意子串是否为回文串。
- 计算从  $s[i]$  开始的最长回文子串的长度。
- 计算以  $s[i]$  结尾的最长回文子串的长度。

Z 函数和 Manacher 算法都会用到类似 Z-box 的概念, 在学习时, 可以对比体会。

- 5. 最长回文子串 模板题
- 647. 回文子串 计算回文子串个数

收藏次数 445  
参与人数 53  
浏览次数 25.9k

## 相关标签

字符串 字符串匹配 后缀数组  
C++ Java Python3 周赛  
数据结构 数据结构与算法 竞赛  
双周赛 滚动哈希

## LeetBook 推荐

- 七章刷完数据结构  
字符串 C++  
38,945 人已读
- 图解算法数据结构  
C++ Java  
310,493 人已读
- 设计数据结构  
数据结构  
12,953 人已读

## 相关讨论

- [力扣刷题攻略] Re: 从零...
- 2024 Rewind | 开启你的年...
- 一鼓作气 | 2024 扣得精彩...
- 「力扣编辑器」使用说明
- 华为鸿蒙 HarmonyOS Next...



力扣 (LeetCode) App  
随时随地参与讨论  
App 扫一扫 即刻参与

- [214. 最短回文串](#)
- [3327. 判断 DFS 字符串是否是回文串](#) ~2300
- [1745. 分割回文串 IV](#) 做到  $\mathcal{O}(n)$
- [1960. 两个回文子字符串长度的最大乘积](#) 2691

## 四、字符串哈希

本题单的大多数题目都可以用字符串哈希解决。

推荐先把 [2156. 查找给定哈希值的子串](#) 做了，对理解多项式哈希的计算方法有帮助。

模板代码见 [我的题解](#)，包含单模哈希和双模哈希。

- [28. 找出字符串中第一个匹配项的下标](#)
- [187. 重复的 DNA 序列](#)
- [1316. 不同的循环子字符串](#) 1837
- [1297. 子串的最大出现次数](#) 做到  $\mathcal{O}(n)$
- [2261. 含最多 K 个可整除元素的子数组](#) 做到  $\mathcal{O}(n^2)$
- [3213. 最小代价构造字符串](#) 2171
- [1044. 最长重复子串](#) 2429
- [718. 最长重复子数组](#)
- [1923. 最长公共子路径](#) 2661
- [3292. 形成目标字符串需要的最少字符串数 II](#) 2662
- [2168. 每个数字的频率都相同的独特子字符串的数量](#)（会员题）同 [2261 题](#)
- [1554. 只有一个不同字符的字符串](#)（会员题）
- [1062. 最长重复子串](#)（会员题）同 [1044 题](#)

## 五、最小表示法

定义**循环左移**操作：把字符串  $s$  的第一个字符  $s[0]$  移除，添加到  $s$  的末尾。例如  $abcd$  操作一次后得到  $bcd a$ 。

问题：你可以执行任意次循环左移操作，计算你能得到的字典序最小的字符串。

推荐先完成 [1163. 按字典序排在最后的子串](#)，再做下面这题。

- [899. 有序队列](#)  $k = 1$  的情况即为最小表示法，请用  $\mathcal{O}(n)$  时间解决

## 六、字典树

- 见 [数据结构题单](#) 第六章。

## 七、AC 自动机

AC 自动机 = 字典树 + KMP。

由于这些题目也可以用其他算法（字符串哈希等）解决，难度分仅供参考。

- [1032. 字符流](#) 1970 **模板题**
- [面试题 17.17. 多次搜索](#) **模板题**
- [1408. 数组中的字符串匹配](#) 做到线性时间复杂度
- [3213. 最小代价构造字符串](#) 2171
- [3292. 形成目标字符串需要的最少字符串数 II](#) 2662

## 八、后缀数组/后缀自动机

由于这些题目也可以用其他算法（字符串哈希等）解决，难度分仅供参考。

- [1163. 按字典序排在最后的子串](#) 1864
- [1754. 构造字典序最大的合并字符串](#) 可以做到  $\mathcal{O}(n + m)$
- [2904. 最短且字典序最小的美丽子字符串](#) 可以做到  $\mathcal{O}(n \log n)$
- [3213. 最小代价构造字符串](#) 2171
- [1044. 最长重复子串](#) 2429
- [718. 最长重复子数组](#)
- [1923. 最长公共子路径](#) 2661
- [1408. 数组中的字符串匹配](#)
- [3076. 数组中的最短非公共子字符串](#)
- [1316. 不同的循环子字符串](#)
- [3388. 统计数组中的美丽分割](#) 做到  $\mathcal{O}(n \log n)$  或  $\mathcal{O}(n)$
- [2564. 子字符串异或查询](#) 见我题解下的评论
- [1698. 字符串的不同子字符串个数](#)（会员题）
- [1062. 最长重复子串](#)（会员题）同 [1044 题](#)
- [3135. 通过添加或删除结尾字符来同化字符串](#)（会员题）

## 九、子序列自动机

上面都是和子串相关的算法，本节是和子序列相关的算法：子序列自动机。

虽然名字有些高大上，但实际上只是预处理  $s[i]$  右侧最近字母  $c$  的下标而已。

- [792. 匹配子序列的单词数](#) 1695 **模板题**
- [514. 自由之路](#) ~2400 做到  $\mathcal{O}(nm)$
- [2014. 重复 K 次的最长子序列](#) 2558
- [1055. 形成字符串的最短路径](#)（会员题）
- [727. 最小窗口子序列](#)（会员题）

关联题单

- [数据结构题单](#) 中的「六、字典树 (trie)」。
- [滑动窗口与双指针题单](#) 中的「§4.2 判断子序列」。
- [贪心题单](#) 中的「三、字符串贪心」。

分类题单

如何科学刷题？

1. 滑动窗口与双指针 (定长/不定长/单序列/双序列/三指针/分组循环)
2. 二分算法 (二分答案/最小化最大值/最大化最小值/第K小)
3. 单调栈 (基础/矩形面积/贡献法/最小字典序)
4. 网格图 (DFS/BFS/综合应用)
5. 位运算 (基础/性质/拆位/试填/恒等式/思维)
6. 图论算法 (DFS/BFS/拓扑排序/最短路/最小生成树/二分图/基环树/欧拉路径)
7. 动态规划 (入门/背包/状态机/划分/区间/状压/数位/数据结构优化/树形/博弈/概率期望)
8. 常用数据结构 (前缀和/差分/栈/队列/堆/字典树/并查集/树状数组/线段树)
9. 数学算法 (数论/组合/概率期望/博弈/计算几何/随机算法)
10. 贪心与思维 (基本贪心策略/反悔/区间/字典序/数学/思维/脑筋急转弯/构造)
11. 链表、二叉树与回溯 (前后指针/快慢指针/DFS/BFS/直径/LCA/一般树)
12. 字符串 (KMP/Z函数/Manacher/字符串哈希/AC自动机/后缀数组/子序列自动机)

我的题解精选 (已分类)

欢迎关注 [B站@灵茶山艾府](#)

如果你发现有题目可以补充进来，欢迎评论反馈。

👍 205 | ☆ 收藏 | ↻ 分享 | ...

↩ 回复讨论

🔔 接收动态

共 53 个回复

最热 ⌵

 灵茶山艾府 · 楼主

来自浙江 · 2024-10-07

题单整理得差不多啦，后面会更新 [HOT 100](#) 的题解，目前 (10.8) 已完成 70%.

大家如果还有什么想看的，可以在评论区留言~

👍 90 | 🔍 回复 6 | ☆ 收藏 | ↻ 分享 | ...

✍ 添加回复

 一堵墙一扇窗

来自湖北 · 2024-10-07

伟大无需多言

👍 21 | 🔍 回复 | ☆ 收藏 | ↻ 分享 | ...

✍ 添加回复

 Raphael

来自广东 · 2024-10-07


《《论灵神为什么是神》》

👍 14 | 🔍 回复 | ☆ 收藏 | ↻ 分享 | ...

✍ 添加回复

 校赛干碎算法梦

来自广东 · 2024-10-15

 题单整理得差不多啦，后面会更新 [HOT 100](#) 的题解，目前 (10.8) 已完成 70%. 大家如果还有什么想看的，可以在评...

灵神能不能再更新下倍增和st算法

👍 10 | 🔍 回复 1 | ☆ 收藏 | ↻ 分享 | ...

✍ 添加回复

 byte

来自天津 · 2024-10-17

刷题的速度没有灵神更新题单的速度快😁

👍 6 | 🔍 回复 | ☆ 收藏 | ↻ 分享 | ...

✍ 添加回复

 Frosty Boydwy7

来自广东 · 2024-10-07

来了来了！题越做越多！

👍 6 | 🔍 回复 1 | ☆ 收藏 | ↻ 分享 | ...

✍ 添加回复

 nlv

来自山西 · 2024-10-07

生熟三学

👍 5 | 🗨 回复 ☆ 收藏 🔄 分享 ...

✎ 添加回复

 站灵-低配版灵神

来自江苏 · 2024-11-09

 题单整理得差不多啦，后面会更新 HOT 100 的题解，目前 (10.8) 已完成 70%。大家如果还有什么想看的，可以在评...

灵神，hot100我正好要二刷，快快更新，还是你的答案比较好背

👍 2 | 🗨 回复 ☆ 收藏 🔄 分享 ...

✎ 添加回复


 有人吗 🤔

来自浙江 · 2024-10-07

两道数据量很小，但要过1e5数据量得上后缀数组的题  
#2904. 最短且字典序最小的美丽子字符串  
#1754. 构造字典序最大的合并字符串

👍 2 | 🗨 回复 1 ☆ 收藏 🔄 分享 ...

✎ 添加回复

 Frosty Boydwy7 🌟

来自广东 · 2024-10-07

 来了来了！题越做越多！

刚好在看 214 🤖

👍 2 | 🗨 回复 ☆ 收藏 🔄 分享 ...

✎ 添加回复

力扣 LeetCode

竞赛

LeetBook

讨论社区

求职

Plus 会员

周边商城

企业服务

在线面试

企业测评

招聘

培训

解决方案

商务

社区合作

活动

赞助竞赛

产品推广

校园合作

关于我们

价值观

工作机会

🔔 📧 知 今

© 2024 领扣网络（上海）有限公司

商务咨询 | 侵权投诉 | 问题反馈 | 加入我们 | 使用条款 | 隐私政策

沪ICP备18019787号-20  沪公网安备31010702007420号 沪ICP证B2-20180578 LeetCode力扣-认证中心  
上海市互联网违法和不良信息举报中心 中国互联网违法和不良信息举报中心