1. 题意

1.1 题目描述

给定一个长度为 n 的字符串 $s[1:n]$。有 q 次询问，每次询问给定两个参数 i, r。你需要求出有多少 l，满足如下条件:

- $1 \le l \le r$。
- $s[i:i+l-1]$ 字典序小于 $s[i+l:i+2l-1]$ 。

$n, q \le 5*10^5$ 。

1.2 数据范围

对于所有数据，$1 \le n, q \le 5*10^5$，$1 \le i+2r-1 \le n$，字符串 s 仅包含小写字母。

1.3 子任务

子任务 1 (20%) : $n, q \le 5*10^3$ 。

子任务 2 (10%) : $n, q \le 10^5$ ，保证 s 中每个字符在 a,b 中随机生成。

子任务 3 (20%) : $n, q \le 10^5$ 。

子任务 4 (20%) : $n, q \le 3*10^5$ 。

子任务 5 (30%) : 无特殊限制。

---

2. 解题过程

2.1 初步转化

对串 s 进行后缀排序，记 rk_i 是 $s[i:n]$ 的排名。

对于一次询问，我们先计算出有多少个 $1 \le l \le r$ 满足 $rk_i < rk_{i+l}$ ，这是二维数点，容易用 BIT 解决，接下来我们只需要计算有多少 l 满足 $s[i:i+l-1] = s[i+l:i+2l-1]$ 且 $rk_i < rk_{i+l}$ ，下面我们称这样的 (i, l) 是合法的。

2.2 对平方串的刻画

处理平方串的问题，我们尝试使用 runs。

下面我们定义一个 run 为三元组 (i,j,p) 满足 $2p \le j-i+1$ 且 p 是 $s[i:j]$ 的最小周期，且 $s_{i-1} \ne s_{i+p-1}, s_{j+1} \ne s_{j+1-p}$ (不妨令 $s_0 = s_{n+1} = -1$)。

我们需要知道，run 和平方串有怎样的关系。可以发现对于 run (i,j,p) ，令 $t = \lfloor \frac{j-i+1}{2p} \rfloor$ ，则对于 $1 \le k \le t, i \le l \le j-2kp+1$ ，$s[l:l+2kp-1]$ 是平方串。

那 s 中每个平方串能通过这样的方式被唯一算到吗？我们尝试证明这个事情。

定理 1：若 $s[l:r]$ 有周期 q ，$s[L:R]$ 有周期 p 且 $q|p$ ，$r-l+1 \ge p$ ，则 $s[L:R]$ 有周期 q 。

原因是，若 $s[l+1:r]$ 存在周期 q ，$s[l:r]$ 存在周期 p 且 $r-l \ge p$ ，那根据 $s[l+1:r]$ 存在周期 q ，我们有 $s_{l+q} = s_{l+2q} = \cdots = s_{l+p}$ ，而根据 $s[l:r]$ 存在周期 p ，我们有 $s_l = s_{l+p}$ 。联立可得 $s_l = s_{l+q}$ ，于是 $s[l:r]$ 存在周期 q 。

同理，若 $s[l:r-1]$ 存在周期 q ，$s[l:r]$ 存在周期 p 且 $r-l \ge p$ ，则 $s[l:r]$ 存在周期 q 。

所以由 $s[l:r]$ 有周期 q ，可以推得 $s[l-1:r], s[l-2:r] \ldots, s[L:r]$ 有周期 q ，再推得 $s[L:r], s[L:r+1], \ldots, s[L:R]$ 有周期 q 。

定理 2：考虑 run (i,j,p) ，对于 $k>0, i \le l \le j-2kp+1$ ，$s[l:l+2kp-1]$ 的最小周期一定是 p 。

考虑反证。首先由定义 p 是 $s[l:l+2kp-1]$ 的周期，若其最小周期 $q<p$ ，根据 Weak Periodicity Lemma，由于 $q+p \le 2kp$ ，所以 $gcd(q,p)$ 也是周期，于是 $gcd(q,p)=q$ ，即 $q|p$ 。

此时 $[l,l+2kp-1]$ 有周期 q ，$[L,R]$ 有周期 p ，$2kp \ge p$ ，那根据定理 1 ，我们有 $[L,R]$ 有周期 q ，与 run 的定义矛盾！得证。

定理 3：对于两个周期相同的 run $(l_1,r_1,p),(l_2,r_2,p)$，其交一定 $<p$ 。

考虑反证。首先二者不可能是包含关系，否则与定义矛盾。现在不妨令 $l_1<l_2<r_1<r_2$ 。由于 $r_1-p+1 \ge l_2$ ，我们有 $s_{r_1+1} = s_{r_1+1-p}$ ，矛盾！

了解了这些定理后就可以证明为什么能唯一算到平方串了：

对于平方串 $s[l:r]$ ，找到其最小周期 p 。则算到 $s[l:r]$ 的 run 周期一定是 p ，否则与定理 2 矛盾；接下来令 $x=l, y=r$ ，若 $s[y+1]=s[y+1-p]$ 则 $y:=y+1$ ，x 同理。这样得到的 run (x,y,p) 一定能算到 $s[l:r]$ ；另一方面，若两个 run $(l_1,r_1,p),(l_2,r_2,p)$ 同时算到 $s[l:r]$ ，则其交 $\ge r-l+1 \ge p$ ，与定理 3 矛盾！

我们就完成了证明部分。

接下来回到问题。我们有办法找到所有平方串 $s[i, i+2l-1]$ 了，但是怎么找出其中 $rk_i < rk_{i+l}$ 的呢？其实很简单：对于 run (i, j, p)，我们只需要看 s_{j+1-p} 是否小于 s_{j+1} 就好了。

原因考虑其中的某个平方串 $s[l : l+2kp-1]$，现在我们要比较 $s[l : n]$ 和 $s[l+kp, n]$，而根据周期我们有 $s_l = s_{l+kp}, s_{l+1} = s_{l+1+kp} \dots, s_{j-kp} = s_j$。现在我们要比较 s_{j-kp+1} 和 s_{j+1}，注意到 $s_{j-kp+1} = s_{j-(k-1)p+1} = \cdots = s_{j-p+1}$，于是就转化成比较 s_{j-p+1} 和 s_{j+1} 了。

2.3 二维数点

最后我们得到做法：对于 run (i, j, p)，若 $s_{j-p+1} < s_{j+1}$，则令 $t = \lfloor \frac{j-i+1}{2p} \rfloor$，则对于 $1 \le k \le t, i \le l \le j-2kp+1$，$(l, kp)$ 是合法的，且每个合法点对只会被考虑到一次。

根据 runs 的性质，所有 run 的 t 之和是 $O(n)$ 的。于是把合法的点对画到平面上，相当于 $O(n)$ 次对一行的点权值 +1，每次查询就是查一列的权值和。这容易离线后用 BIT 实现。

总复杂度 $O((n+q)\log n)$。

3. 命题总结

本题是 [NOI2023] 字符串 的改编题，主要考察选手对 后缀数组和 runs 的掌握程度，思维难度简单，代码难度中等。

熟悉 runs 的选手能轻松得到全部分数。不熟悉 runs 的选手通过其他字符串相关的算法也能得到较高的分数。希望能有更多选手掌握 runs 这一方便的工具。

此外，出题人不知道是否存在不需要 runs 的做法能达到相同的复杂度。

4. 参考资料

[NOI2023] 字符串 及相关题解。

command_block《Lyndon & Runs》，https://www.luogu.com/article/d4y3zqqv