2 Three Indices

2.1 题目来源

Petrozavodsk Winter 2019. Day 5. Gennady Korotkevich Contest 4 Problem C[2].

2.2 题目大意

称一个字符串 t 是关于字符串 w 的光滑变换，当且仅当存在一个整数 $m \ge 1$ 和一个字符串序列 $w_0, w_1, \ldots, w_m$，满足：

- $w_0 = w$。
- $\forall 0 < i \le m, |w_i| = |w|$。
- $\forall 0 < i \le m$，w_i 与 w_{i-1} 至多有一个位置的字符不同。
- $t = w_0 + w_1 + \ldots + w_m$，其中 $+$ 表示字符串拼接。

给定一个长度为 n 的字符串 $s = s_1 s_2 \ldots s_n$。求有多少个三元组 (i, j, k) 满足：

- $1 \le i < j < k \le n$。
- $s_{i\ldots k}$ 是关于 $s_{i\ldots j}$ 的光滑变换。

2.3 数据范围

对于所有数据，满足 $4 \le n \le 10^5$，字符集为全部小写英文字母。

2.4 解题过程

对于一个字符串 s，定义 $s_{[l,r]}$ 表示 s 中第 l 个字符到第 r 个字符构成的子串。
对于两个长度分别为 n, m 的字符串 S, T，定义：

- $\mathrm{LCP}(S, T)$ 表示 S 和 T 的最长公共前缀长度，即最大的 i 满足 $S_{[1,i]} = T_{[1,i]}$。
- $\mathrm{LCS}(S, T)$ 表示 S 和 T 的最长公共后缀长度，即最大的 i 满足 $S_{[n-i+1,n]} = T_{[m-i+1,m]}$。
- $\mathrm{LCP}'(S, T)$ 表示 S 和 T 的最长单差异前缀长度，即最大的 i 满足 $S_{[1,i]}$ 与 $T_{[1,i]}$ 至多只有一个对应位置的字符不同。

[2]https://qoj.ac/contest/2102/problem/12313

- $\mathrm{LCS}'(S,T)$ 表示 S 和 T 的最长单差异后缀长度，即最大的 i 满足 $S_{[n-i+1,n]}$ 与 $T_{[m-i+1,m]}$ 至多只有一个对应位置的字符不同。

称 $|w|$ 为循环长度。考虑枚举循环长度为 L，统计有多少个 $s_{[l,r]}$ 子串具有 L 的循环长度。

对于任意一个合法的拥有 L 循环长度的串 $s_{[l,r]}$，其所有分割点 $\bmod L$ 的值应当全部相等。将整个串按照每段 L 个字符分段，对于一个具有 L 循环长度的串 $s_{[l,r]}$，分析其分割后的相邻两个子段 w_i, w_{i+1} 在原串中的位置：

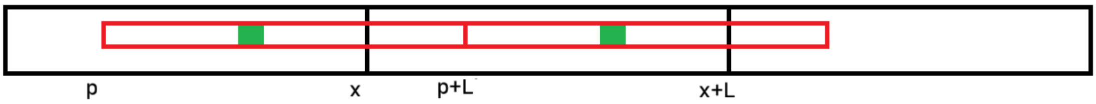

图 1: w_i, w_{i+1} 与原串划分的位置关系

图 1 中黑色块为原串按照 L 个字符一块进行划分所得的块，两个红色块分别为 w_i 与 w_{i+1}，绿色对应的位置是 w_i 与 w_{i+1} 中不同字符的位置。

由于除了绿色位置外，其余位置两个红色块均对应相等，因此 $\mathrm{LCS}'(s_{[1:x]}, s_{[1,x+L]}) \geq x-p$，$\mathrm{LCP}(s_{[x+1,n],s_{x+L+1,n}}) \geq p+L-x$。同时，容易发现在满足上述两个条件时，$w_i$ 和 w_{i+1} 必然至多只相差一个字符。

当 w_i 与 w_{i+1} 不同的字符在原串的分割点 x 右侧时，相对位置关系如图二所示：

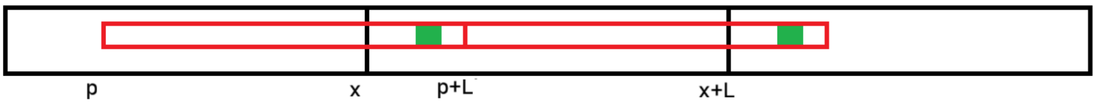

图 2: 不同位置在 x 右侧的情况

此时有 $\mathrm{LCS}(s_{[1:x]}, s_{[1,x+L]}) \geq x-p$，$\mathrm{LCP}'(s_{[x+1,n],s_{x+L+1,n}}) \geq p+L-x$。

总和上述两个条件，“两者之一成立”与“w_i 和 w_{i+1} 至多有一个字符不同”两个条件等价。因此，枚举分割点 $x = L, 2L, \ldots$，求解出 $\mathrm{LCP}(s_{[x+1,n]}, s_{[x+L+1,n]})$，$\mathrm{LCS}(s_{[1,x]}, s_{[1,x+L]})$，$\mathrm{LCP}'(s_{[x+1,n]}, s_{[x+L+1,n]})$，$\mathrm{LCS}'(s_{[1,x]}, s_{[1,x+L]})$ 四个变量后，可以解出 p 的合法范围为：

$$[x-\mathrm{LCS}'(s_{[1:x]}, s_{[1,x+L]}), \mathrm{LCP}(s_{[x+1,n],s_{x+L+1,n}})+x-L] \cup [x-\mathrm{LCS}(s_{[1:x]}, s_{[1,x+L]}), \mathrm{LCP}'(s_{[x+1,n],s_{x+L+1,n}})+x-L]$$

同时容易解出 $p \bmod L$ 的合法范围。该范围为至多两个区间。

问题转化为：求解有多少对 (p,l,r) 满足 $0 \leq p < L$，$1 \leq l \leq r \leq \frac{n}{x}$，且 x 在 $l,\ldots,r$ 这些限制上全部合法。

该问题可以使用扫描线与线段树来解决。从前往后扫描每个限制，使用线段树维护所有 $0 \leq p < L$ 的当前连续合法后缀长度。每次扫描到一个限制时，将所有合法限制外的区间

赋值为 0，将所有合法限制内的区间做区间 +1 操作，并将修改后线段树上维护的所有数的和加入到答案中即可。

求解 LCP, LCS 可以对 s 的正反串建立后缀数组，同时使用 ST 表预处理 Height 数组的区间最小值，做到 $O(n\log n)$ 预处理，$O(1)$ 查询。求解 LCP′, LCS′ 只需要先求解一遍 LCP 或 LCS，可得到第一个不相等的位置，跳过这个位置再求解一次 LCP 或 LCS 即可，复杂度仍然为 $O(1)$。

对于枚举的一个循环长度 L，需要枚举 $\frac{n}{L}$ 个分割位置 x，每次需要求解 4 次 LCP, LCS 与 LCP′, LCS′，随后进行 $O(1)$ 次线段树区间赋值为 0 和区间加法操作，单次复杂度 $O(\log L)$。

根据调和级数分析，分割点总数为 $\sum_{1\le L\le n}\lfloor\frac{n}{L}\rfloor = O(n\log n)$。至此，该问题在 $O(n\log^2 n)$ 的复杂度内被解决。

2.5 参考资料

无。