

string

- ▶ 对于一个串 T ，将其划分为若干连续子串，使得至少存在一个串出现了两次，且所有串中本质不同的个数最少。将这个最小值称为 T 的价值。
- ▶ 给定串 S ， m 次询问，每次询问某个连续子串的价值。无解输出 -1 。
- ▶ Source: CF1043G。

正解

- ▶ 分类讨论。
- ▶ 如果所有字母都只出现过一次，答案是 -1 。
- ▶ 否则，肯定有至少一个字母出现过两次，那么可以找到两次出现的位置在两侧切开，所以答案肯定 ≤ 4 。
- ▶ 答案为 1 时串需要是循环串。
- ▶ 答案为 2 时串需要形如 AAB ， ABA 或 BAA 。
- ▶ 答案为 3 时串需要形如 $aBaC$ ， $BAAC$ 或 $BaCa$ 。
- ▶ 其它情况答案为 4。
- ▶ 这里的大写字母表示一个非空字符串，小写字母表示一个字母。

正解

- ▶ 建出原串和反串的SA。
- ▶ 循环串可以直接枚举循环节判定。一个优化是考虑只枚举长度的质因子作为循环次数。
- ▶ 判定是否是一个AA串的前缀/后缀或者是否包含一个AA串都可以通过预处理数组 b_i, c_i 分别表示以 i 为左/右端点的最小的AA串的长度来实现。预处理直接使用NOI2016优秀的拆分的方法实现。
- ▶ 判定 $aBaC$ 和 $BaCa$ 只需预处理每个字母的前驱后继。
- ▶ 判定ABA可以直接考虑使用区间求border的方法，但由于此题只需要判定，我们有一个简洁的 $O(\sqrt{n})$ 的方法来求。

正解

- ▶ 只需求出某个区间的最小border，考虑对于 $< \sqrt{n}$ 的暴力枚举，对于 $\geq \sqrt{n}$ 的可以发现它在sa数组上到左端点 l 的距离一定 $\leq \sqrt{n}$ 。这个性质是由于这个串的出现原串中的位置一定不交，否则可以找出更小的border。
- ▶ 当然你也可以使用基本子串字典来做到更优秀的复杂度。
- ▶ 时间复杂度 $O(m\sqrt{n})$ 或 $O((n+m)\log^2 n)$ 或 $O((n+m)\log n)$ 。