

《往事》解题报告

厦门双十中学 汪文潇

1 试题来源

这题是我的原创题。

2 试题大意

给定一棵字符集大小为300的有 n 个节点的字典树，定义字典树上每个点对应的字符串为其从其到根的简单路径上的字符依次形成的字符串。

对于一对不同的点 u 和 v ，定义它们之间的相似度为他们对应的字符串的最长公共前缀和最长公共后缀的长度之和。

求该树上所有点对相似度的最大值， $n \leq 200000$ 。

3 算法介绍

先考虑一个简化后的问题，即假设该字典树是一条链的情况。

此时可以将其看成一个字符串，询问该字符串所有后缀中，最长公共前缀和最长公共后缀长度和最大的一对的长度。

注意到，最长公共前缀和最长公共后缀其实是相对独立的，那么我们不妨枚举其中之一。假定我们枚举了最长公共前缀的长度，那么我们要求的其实就是保证了公共前缀的长度后，最长的公共后缀的长度。

注意到公共前缀的性质具有传递性，即A和B有长度为 k 的公共前缀，B和C有长度为 k 的公共前缀，那么A和C就有长度为 k 的公共前缀。结合后缀数组的性质，不难想出这样一个做法：从大到小枚举公共前缀的长度 len ，并将所有后缀分成若干集合，使得每个集合内任意两个后缀的最长公共前缀长度都不少于 len 。在 len 减少的过程中，将集合间进行合并，并维护每个集合内最长公共后缀长度最长的一对后缀即可。

维护每个集合的信息可以通过另一个后缀数组以及线段树合并或启发式合并来实现。

考虑如何将其扩展到树上：显然只要对字典树实现一个后缀数组即可，相应的，在维护每个集合时，也需要使用线段树合并或者启发式合并的技巧。

该算法的时间复杂度为 $O(n \log n)$ 。