

第十二次课 (字符串)

试题册

软件协会

题目列表

游戏人生第二季定档

字符串匹配

随机生成的字符串岂不是随便搞都能过

字符串模糊匹配

最长公共子串

最短周期

前缀出现次数

前缀出现次数2

字典树

Browser Games

最长异或路径

最长回文子串

重要提示 Java 和 Python 程序在题目所标时间空间限制基础上有 2 秒的额外运行时间和 64 MB 的额外空间

A. 游戏人生第二季定档

1 秒, 256 MB



记字符串 S 为: never gonna give you up, never gonna let you down, never gonna run around and desert you, never gonna make you cry, never gonna say goodbye, never gonna tell a lie and hurt you

去除标点符号后，按空格拆分串 S 得到单词集合 T ， T 的元素大致如上图所示(但 A LIE 是两个元素)

第十二次课考勤题目 (っω´)っ”

我们建议你心算本题而不是写代码，若你对下文概念有不明白的，请翻阅课件 OvO

输入

无输入

输出

第一行一个整数，代表 $|S|$

第二行一个字符串，代表 T 里字典序第二小的串

第三行若干个字符串，代表 T 中所有回文串

第四行若干个字符串，代表 T 中所有存在真前缀等于真后缀的字符串

第三第四行若有多个结果，按字典序输出；若没有任何结果输出空行

B. 字符串匹配

10 秒, 40 MB

给定由不含空格的可显示 ASCII 字符组成的字符串 S, T ，记下标从 1 开始，请你求出 S 中出现 T 的每个位置的下标。若一次也不出现，输出 -1 。

输入

输入两行，依次是字符串 $S, T(1 \leq |S|, |T| \leq 2 \times 10^6)$

输出

输出若干整数，代表出现位置的下标。整数间用一个空格隔开。

样例

样例输入 1	样例输出 1
nunhehhehaaaaaaaaaaaaaaaaaa aaaaaaaa	10 11 12 13 14 15 16 17 18 19 20
样例输入 2	样例输出 2
ababababa ababa	1 3 5
样例输入 3	样例输出 3
ParsleySageRosmaryAndThyme Chth0lly	-1

说明

请使用较为快速的读入和输出方式。

C. 随机生成的字符串岂不是随便搞都能过

1 秒, 512 MB

给定字符串 a, b , 给出下面 C++ 程序期望的运行结果。

```
#include <iostream>

long long cnt = 0;

int main()
{
    std::string a, b;
    std::cin >> a >> b;
    for (size_t i = 1; i <= a.size(); i++)
    {
        for (size_t j = 1; j <= a.size() && i + j <= b.size(); j++)
        {
            cnt += (a.substr(0, i) + a.substr(0, j) == b.substr(0, i + j));
        }
    }
    std::cout << cnt << std::endl;
}
```

输入

第一行一个字符串 a , 第二行一个字符串 b 。

$1 \leq |a|, |b| \leq 10^5$, 只包含小写字母。

输出

输出一行一个整数, 表示程序的运行结果。

样例

样例输入 1	样例输出 1
aab aaa	3

说明

数据可能有点弱, 不要水过去就完事就行 :-)

D. 字符串模糊匹配

2 秒, 128 MB

给定系数 k , 定义字符串 A, B 模糊匹配当且仅当 $|A| = |B|$ 且最多有 k 个位置字符不同。

给定由不含空格的可显示 ASCII 字符组成的字符串 S, T , 记下标从 1 开始, 请你求出 S 的模糊匹配成功 T 的每个下标。若一次也不出现, 输出 -1 。

输入

首先输入一行一个整数 $k(0 \leq k \leq 5)$

接着输入两行, 依次是字符串 $S, T(1 \leq |S|, |T| \leq 10^6)$

输出

输出若干整数, 代表出现位置的下标。整数间用一个空格隔开。

样例

样例输入 1	样例输出 1
3 nunhehhehaaaaaaaaaaaaaaaaaa aaaaaaaa	7 8 9 10 11 12 13 14 15 16 17 18 19 20

样例输入 2	样例输出 2
2 acacababa ababa	1 3 5

样例输入 3	样例输出 3
5 ParsleySageRosmaryAndThyme Chth0lly	-1

说明

请使用较为快速的读入和输出方式。

E. 最长公共子串

1 秒, 128 MB

给定 n 个由不含空格的可显示 ASCII 字符组成的字符串，第 i 个字符串为 S_i ，求它们的最长公共子串的长度。

输入

输入一行一个整数 $n(2 \leq n \leq 10^6)$

接下来输入 n 行，每行一个字符串 $S_i(1 \leq |S_i| \leq 10^6)$

保证 $\sum_{i=1}^n |S_i| \leq 10^6$

输出

输出一行一个整数，代表最长公共子串的长度

样例

样例输入 1	样例输出 1
4 guodongsensei senseiguodong hentaiguodong guodongdongguo	7

样例输入 2	样例输出 2
3 abcd ABC 01234	0

样例输入 3	样例输出 3
6 smile sweet sister sadistic suprice suanfa	1

说明

请使用较为快速的读入和输出方式。

F. 最短周期

1 秒, 128 MB

若对字符串 S , $\exists p \in N, \forall i \in [1, |S| - p]$ 都有 $S[i] = S[i + p]$, 那么 p 是 S 的周期。请你求出字符串 S 的最小周期

输入

输入一行一个由不含空格的可显示 ASCII 字符组成的字符串 $S(1 \leq |S| \leq 2 \times 10^6)$

输出

输出一行一个整数，代表 S 的最小周期。

样例

样例输入 1	样例输出 1
gggggg	1

样例输入 2	样例输出 2
buyaohuida!buyaohuida!buyaohuida!	11

样例输入 3	样例输出 3
nunhehhehaaaaaaaaaaaaaaaaaa	27

说明

如果一个字符串不能分成相等的两部分或多部分，那么周期是串长。

G. 前缀出现次数

1 秒, 128 MB

给定字符串 S ，求出 S 的每个真前缀在 S 的出现次数。为了防止输出过长，设长为 i 的前缀出现次数为 $t(i)$ ，你只需要输出 $\sum_{i=1}^{|S|-1} i \times t(i)$ 即可

输入

输入一行一个由不含空格的可显示 ASCII 字符组成的字符串 $S(2 \leq |S| \leq 2 \times 10^6)$

输出

输出一行一个整数，代表 $\sum_{i=1}^{|S|-1} i \times t(i)$

样例

样例输入 1	样例输出 1
ababc	13

样例输入 2	样例输出 2
abacaba	29

样例输入 3	样例输出 3
1r580	10

说明

对 `ababc`，有 $t(1) = t(2) = 2, t(3) = t(4) = 1$

对 `abacaba`，有 $t(1) = 4, t(2) = t(3) = 2, t(4) = t(5) = t(6) = 1$

对 `1r580`，有 $t(1) = t(2) = t(3) = t(4) = 1$

保证输出结果不会大于 `long long` 所能表示的范围

H. 前缀出现次数2

1 秒, 128 MB

给定字符串 S, T , 求出 S 的每个前缀在 T 的出现次数。为了防止输出过长, 设长为 i 的前缀出现次数为 $t(i)$, 你只需要输出 $\sum_{i=1}^{|S|} i \times t(i)$ 即可

输入

输入两行, 分别是只由小写字母组成的字符串 $S, T(1 \leq |S|, |T| \leq 2 \times 10^6)$

输出

输出一行一个整数, 代表 $\sum_{i=1}^{|S|} i \times t(i)$

样例

样例输入 1	样例输出 1
aaa aaaaa	22

样例输入 2	样例输出 2
kmp kkkmpmmpkm	11

样例输入 3	样例输出 3
guodong baicha	0

说明

对第一个样例, $t(1) = 5, t(2) = 4, t(3) = 3$

对第二个样例, $t(1) = 4, t(2) = 2, t(3) = 1$

对第三个样例, $\forall 1 \leq i \leq 7, t(i) = 0$

保证输出结果不会大于 `long long` 所能表示的范围

I. 字典树

1 秒, 128 MB

给定 n 个只由小写字母组成字符串，第 i 个字符串为 S_i 。接下来有 m 个询问，第 i 个询问给定一个只由小写字母组成的字符串 T_i ，对于每个询问，求有多少个字符串 S 满足其前缀为 T_i 。不保证 S 各不相同。

输入

首先输入一行一个整数 $n(1 \leq n \leq 10^5)$

接下来输入 n 行，第 i 行一个字符串 $S_i(1 \leq S_i \leq 5 \times 10^5)$ 。保证 $\sum_{i=1}^n S_i \leq 5 \times 10^5$

接下来输入一行一个整数 $m(1 \leq m \leq 10^5)$

接下来输入 m 行，第 i 行一个字符串 $T_i(1 \leq T_i \leq 5 \times 10^5)$ 。保证 $\sum_{i=1}^n T_i \leq 2 \times 10^6$

输出

对于每个询问，输出一行一个整数，代表询问结果

样例

样例输入 1	样例输出 1
8 smile sweet sister sadistic suprice service space suanfa 6 string s su sum sad t	0 8 2 0 1 0

样例输入 2	样例输出 2
7 acautomaton acautomaton acautomation suffixarray suffixautomation generalsuffixautomation sunday 7 trie acautomaton acautomat acauto su sun sudormrf	0 2 3 3 3 1 0

J. Browser Games

2 秒, 1024 MB

In the upcoming n days, n browser games will be released on a new website. According to the plan, the administrator will release a new game per day. Users have to open the corresponding URL (Uniform Resource Locator) and get feedback from the server to download a game.

However, the setup of the server uses unreadable legacy codes. Once a user somehow finds the URL of an unreleased game, the data of the game would leak out. To temporarily fix the problem, the administrator decided to add a series of **confirmation prefixes**, which are non-empty strings, at the server-side. The server will respond with the correct game data when the requested URL does correspond to a game (no matter released or unreleased) and at least one **confirmation prefix** is a prefix of the URL; otherwise the server will declare that the game is not found.

To make the work easier, the administrator asks you to find the minimum number of **confirmation prefixes** the server required to avoid data leaks every time after a new game release.

输入

The first line contains an integer n ($1 \leq n \leq 5 \times 10^4$), indicating the number of browser games to be released.

In the next n lines, the i -th line contains a non-empty string, consisting of only lowercase letters ('a' to 'z'), dots('.') and forward slashes('/'), indicating the URL of the browser game released on the i -th day,

It is guaranteed that the length of each given URL is at most 50, and no given URL is the prefix of any other given URL.

输出

Output n lines, the i -th of which contains an integer indicating the minimum number of required confirmation prefixes after the i -th new game released.

样例

样例输入 1	样例输出 1
3 ufoipv.ofu hsbocmvfgboubtz.kq hfotijo.njipzp.dpn/kb	1 2 2

K. 最长异或路径

1 秒, 64 MB

给定一棵 n 个点的带权树，结点下标从 1 开始到 n 。寻找树中找两个结点，求最长的异或路径。

异或路径指的是指两个结点之间唯一路径上的所有边权的异或。

输入

第一行一个整数 n ，表示点数。

接下来 $n-1$ 行，给出 u,v,w ，分别表示树上的 u 点和 v 点有连边，边的权值是 w 。

$$1 \leq n \leq 10^5, 0 < u, v \leq n, 0 \leq w < 2^{31}$$

输出

一行，一个整数表示答案。

样例

样例输入 1	样例输出 1
4 1 2 3 2 3 4 2 4 6	7

说明

最长异或序列是 1,2,3，答案是 $7 = 3 \oplus 4$

原题 [P4551](#)，如果您 AC 这里的本题后想要经过强数据检验，请前往洛谷交题

L. 最长回文子串

1 秒, 256 MB

给定只由英文小写字母组成的字符串 S ，求 S 中最长回文子串的长度

输入

输入一行一个字符串 $S(|S| \leq 1.1 \times 10^7)$

输出

输出一行一个整数代表答案

样例

样例输入 1	样例输出 1
aaa	3

说明

强数据请前往洛谷 [p3805](#)