

1131: 基因相关性

题目描述

为了获知基因序列在功能和结构上的相似性，经常需要将几条不同序列的DNA进行比对，以判断该比对的DNA是否具有相关性。

现比对两条长度相同的DNA序列。定义两条DNA序列相同位置的碱基为一个碱基对，如果一个碱基对中的两个碱基相同的话，则称为相同碱基对。接着计算相同碱基对占总碱基对数量的比例，如果该比例大于等于给定阈值时则判定该两条DNA序列是相关的，否则不相关。

输入格式

有三行，第一行是用来判定出两条DNA序列是否相关的阈值，随后2行是两条DNA序列(长度不大于500)。

输出格式

若两条DNA序列相关，则输出“yes”，否则输出“no”。

输入样例

```
0.85
ATCGCCGTAAGTAACGGTTTTAAATAGGCC
ATCGCCGGAAGTAACGGTCTTAAATAGGCC
```

输出样例

```
yes
```

解析

- 0、定义double n, 存储输入的相似度数值。
定义double num, 存储在两个字符串中有多少在相同的位置是相同的字符。
定义int len存储给定字符串的长度
定义string s1存储第一条字符串
定义string s2存储第二条字符串
- 1、读取给定的相似度数值n, 使用cin
- 2、读入给定的两条字符串, 因为不存在空格的问题, 所以可以使用cin, getline.
- 3、因为两个字符串的长度是相同的, 所以我们可以获取任意string的长度, 使用s1.size().
- 4、使用一个for循环, 遍历两个字符串中的每一个字符, 判断两个字符串在相同的索引位置上的值是否相同,
如果相同则将num进行累加。
- 5、求解当前的相似度res, 即使用 num / len , 因为要求解的值是个浮点数。
- 6、如果res > n, 输出yes
否则, 输出no

编码

```
#include <bits/stdc++.h>

using namespace std;

int main(int argc, char **argv) {
    //目标比例, 大于比例输出yes, 小于比例输出no
    double ratio;
    //两个字符串相似的数量
    double num;
    //DNA序列的长度
    int len;
    //读入三个关键数据
    string s1, s2;
    cin >> ratio >> s1 >> s2;
    len = s1.size();
    //遍历每一个字符, 进行比较, 判断二者是否相同
    for (int i = 0; i < len; i++) {
        if (s1[i] == s2[i]) {
            //累加相同的字符数
            num++;
        }
    }
    //计算出当前的比例值
    double res = num / len;
    //与目标比例作比较
    if (res >= ratio) {
        cout << "yes";
    } else {
        cout << "no";
    }
}
```

```
}  
    return 0;  
}
```

逻辑航线培优教育，信息学奥赛培训专家。

扫码添加作者获取更多内容。

