

## 3716 - DNA Regions

#### **Europe - Central - 2006/2007**

A DNA sequence or genetic sequence is a succession of letters representing the primary structure of a real or hypothetical DNA molecule or strand, with the capacity to carry information. The possible letters are A, C, G, and T, representing the four nucleotide subunits of a DNA strand: adenine, cytosine, guanine and thymine bases covalently linked to phospho-backbone.

DNA sequences undergo mutations during the evolution of species, which means that some letters are randomly replaced with others. Therefore, the DNA sequences of two closely related species are very similar, and the difference increases as the distance between the species increases. The mutations do not occur with uniform frequency throughout the sequence; typically there are fewer mutations at the biologically important parts, since even a single mutation can be lethal at such a place. On the other hand, if a part of the sequence does not carry any biologically relevant information, then mutations on this part have no effect. It follows that if we compare the DNA sequences of two species and a particular region of the sequence contains fewer than the average number of mutations, then most probably this part of the sequence plays an important biological role. Therefore, it is of crucial importance to identify such regions. More precisely, a conserved region is a consecutive interval of the DNA sequence such that in this region at most *p* percent of the letters are different in the two sequences. Your task is to write a program that, given two DNA sequences, finds the longest conserved region.

#### Input

The input contains several blocks of test cases. Each case begins with a line containing two integers:  $1 \le n \le 150000$ , the length of the genetic sequences and  $1 \le p \le 99$ , the maximum percentage of mutated

letters allowed in a conserved region. This is followed by two lines, each containing a DNA sequence of length n. The sequence contains only the letters `A', `C', `G', and `T'.

The input is terminated by a test case with n = 0.

#### **Output**

For each test case, you have to output a line containing a single integer: the length of the longest conserved region between the two sequences. If there are no conserved regions in the input, then output `No solution.' (without quotes).

### **Sample Input**

14 25
ACCGGTAACGTGAA
ACTGGATACGTAAA
14 24
ACCGGTAACGTGAA
ACTGGATACGTAAA
8 1
AAAAAAAA
CCCCCCCC
8 33
AAACAAAA
CCCCCCCC

# **Sample Output**

```
8
7
No solution.
1
```

Central 2006-2007