

Projeto Genoma

Um grande projeto mundial está em curso para mapear todo o material genético do ser humano: o Projeto Genoma Humano. As moléculas de DNA (moléculas que contêm material genético) podem ser representadas por cadeias de caracteres que usam um alfabeto de apenas 4 letras: 'A', 'C', 'T' e 'G'. Um exemplo de uma tal cadeia é:

TCATATGCAAATAGCTGCATACCGA

Nesta tarefa você deverá produzir uma ferramenta muito utilizada no projeto Genoma: um programa que procura ocorrências de uma pequena cadeia de DNA (que vamos chamar de p) dentro de uma outra cadeia de DNA (que vamos chamar de t). Você deverá procurar dois tipos de ocorrência: a "direta" e a "complementar invertida".

Uma ocorrência *direta* é quando a cadeia p aparece como subcadeia dentro de t . Por exemplo, se

$p = \text{CATA}$

$t = \text{TCATATGCAAATAGCTGCATACCGA}$,

então p ocorre na forma direta na posição 2 e na posição 18 de t .

Uma ocorrência *complementar invertida* depende da seguinte correspondência entre as letras do DNA: 'A' \Leftrightarrow 'T' e 'G' \Leftrightarrow 'C'. "Complementar o DNA" significa trocar as letras de uma cadeia de DNA seguindo essa correspondência. Se complementarmos a cadeia CATA, vamos obter GTAT. Mas além de complementar, é preciso também inverter, ou seja, de GTAT obter TATG. E é esta cadeia que deverá ser procurada, no caso da ocorrência complementar invertida. Assim, se p e t são as mesmas cadeias do exemplo anterior, então p ocorre na forma complementar invertida na posição 4 de t .

1. Tarefa

Sua tarefa é escrever um programa que, dadas duas cadeias p e t , onde o comprimento de p é menor ou igual ao comprimento de t , procura todas as ocorrências diretas e todas as ocorrências complementares invertidas de p em t .

2. Entrada de Dados

O arquivo GENOMA.IN contém vários conjuntos de teste. Cada conjunto de teste é composto por três linhas. A primeira linha contém dois inteiros positivos, M e N , $M \leq N$, que indicam respectivamente o comprimento das cadeias de DNA p e t , conforme descrito acima. A segunda linha do conjunto de teste contém a cadeia p , e a terceira linha contém a cadeia t , onde p e t são compostas utilizando apenas os caracteres 'A', 'C', 'G' e 'T'. O final do arquivo de testes é indicado quando $M = N = 0$ (este último conjunto de testes não é válido e não deve ser processado).

O arquivo GENOMA.IN contém ao menos um conjunto de teste que deve ser processado.

Exemplo de Entrada

```
2 4
AC
TGGT
4 25
CATA
TCATATGCAAATAGCTGCATACCGA
0 0
```

3. Saída de Dados

Seu programa deve produzir um arquivo de saída chamado GENOMA.OUT. Para cada conjunto de teste do arquivo de entrada seu programa deve produzir quatro linhas no arquivo de saída. A primeira linha deve conter um identificador do conjunto de teste, no formato “Teste n ”, onde n é numerado a partir de 1. Na segunda linha deve aparecer a lista, em ordem crescente, com a posição inicial de cada ocorrência, na forma direta, do padrão p na sequência t . Na terceira linha deve aparecer a lista, em ordem crescente, com a posição inicial de cada ocorrência, na forma complementar invertida, do padrão p na sequência t . A quarta linha deve ser deixada em branco. A grafia mostrada no Exemplo de Saída, abaixo, deve ser seguida rigorosamente.

Exemplo de Saída

```
Teste 1
ocorrendia direta: 0
ocorrendia complementar invertida: 3
```

```
Teste 2
ocorrendia direta: 2 18
ocorrendia complementar invertida: 4
```

(esta saída corresponde ao exemplo de entrada acima)

4. Restrições

$1 \leq M \leq 15000$

$1 \leq N \leq 15000$

$M \leq N$

$M = 0$ e $N = 0$ apenas para indicar o fim do arquivo de entrada