

nikolai

[Fiche publique](#)

[Mon profil](#)

► [Authentification](#)

[Déconnexion](#)

Langue : 

Plan du site

Progresser

- [Présentation](#)
- [Cours et problèmes](#)
- [Questions fréquentes](#)
- [Forum d'entraide](#)

Enseigner

- [Présentation](#)
- [Groupes et classes](#)

Concourir

- [Présentation](#)
- [Classement](#)
- [Épreuves de concours](#)
- [Résultats](#)

Olympiades

- [Présentation](#)
- [Sélection](#)
- [Résultats](#)
- [Autres olympiades](#)

L'association

- [Présentation](#)
- [Historique](#)
- [Stages d'entraînement](#)
- [Nous contacter](#)
- [Signaler une anomalie](#)

797 personnes connectées

Demandes d'aide

[lenaliscoet](#) ²⁰²⁰ [auj.](#)

Cour de récréation

[gryfenfer59](#) ²⁰²² [auj.](#)

Choix des emplacements

[mohamedboughmadi...](#) [auj.](#)

C++

Fragments d'ADN



Sujet

Résoudre

Conseils

Activité

Correction

Depuis que les biologistes disposent de machines capables de séquencer des morceaux de génome, ils récupèrent des giga-octets de données. Cependant des données brutes ne présentent aucun intérêt si l'on ne dispose pas d'algorithmes suffisamment efficaces pour les analyser. Ce sujet vous propose un de ces nombreux problèmes d'analyse de séquence d'ADN.



On commence par vous décrire une grande chaîne d'ADN, c'est-à-dire une suite de lettres parmi C,G,T,A. Ensuite on vous donne des motifs d'ADN correspondant à des gènes, c'est-à-dire une suite de quelques dizaines de lettres (toujours parmi C,G,T,A). Le but est de déterminer pour chaque gène quelle est la zone de la grande chaîne d'ADN qui ressemble le plus au motif de ce gène.

Concrètement, on veut minimiser le nombre de différences entre le code du gène et le code d'un fragment de la grande chaîne d'ADN (le fragment doit être de même longueur que le code du gène et doit être entièrement inclus dans la chaîne d'ADN).

Pour simplifier, on va travailler en considérant que l'ADN est simplement une suite de 0 et de 1 (remarquez que cela ne change pas fondamentalement le problème). On supposera aussi que tous les gènes considérés sont des séquences de même longueur. Ce qu'on vous demande précisément dans ce sujet, c'est de déterminer pour chaque gène quel est le nombre minimum de différences que l'on peut obtenir entre le code du gène et le code d'un fragment de même longueur entièrement inclus dans la grande chaîne d'ADN.

Notez que vous ne pouvez pas retourner les gènes ou la chaîne d'ADN.

Limites de temps et de mémoire (C++)

- Temps : 0,4 s sur une machine à 1 GHz.
- Mémoire : 32 000 ko.

Contraintes

L'espion est démasqué !

[lpricorné](#) ²⁰²¹ auj.

Code secret deux fois

Réponse obtenue auj.

Amis zoinxiens

[alex3er](#) ²⁰¹⁰ hier

Intervention des pompie...

Réponse obtenue 4 j.

Course d'orientation

Réponse obtenue 01/03

Championnat de ping-po...

Réponse obtenue 19/03/19

Fibre optique

Réponse obtenue 08/06/18

Plus long palindrome

Réponse obtenue 08/06/18

Moyenne hexadécimale

Problèmes résolus

[julie7](#) ²⁰²² 12 h45

Tarifs dégressifs

[anne-laure5](#) ²⁰²² 12 h45

Encore des punitions

[lisa-sophiebrosem...](#) 12 h45

Présentation

[marygrf](#) ²⁰²² 12 h45

Deux codes secrets

[titi9](#) ²⁰²⁰ 12 h45

Course avec les enfants

[2bbjachgobi](#) ²⁰²² 12 h45

Table de multiplication

[marygrf](#) ²⁰²² 12 h44

Code secret deux fois

[gsh10601](#) ²⁰²¹ 12 h44

L'éclipse

[lpollini](#) ²⁰²² 12 h44

Bonbons pour tout le mo...

[jarraya-med](#) ²⁰²⁰ 12 h44

Le secret du Goma

[pjurine1](#) ²⁰²⁰ 12 h44

Transport des bagages

[2caflmagaer](#) ²⁰²⁰ 12 h44

Kermesse

- $1 \leq G \leq 100$, où G est le nombre de gènes que l'on cherche à reconnaître approximativement dans la grande chaîne.
- $1 \leq L \leq 30$, où L est la longueur des segments d'ADN des gènes considérés.
- $L \leq N \leq 100\,000$, où N est la longueur de la grande chaîne d'ADN à étudier.

Entrée

- La première ligne de l'entrée contient trois entiers séparés par des espaces : G , L et N .
- Chacune des G lignes suivantes contient L entiers (0 ou 1) séparés par des espaces décrivant un gène.
- La dernière ligne contient N entiers (0 ou 1) séparés par des espaces décrivant la grande chaîne.

Sortie

Vous devez écrire G lignes contenant chacune un entier. La ligne $i^{\text{ème}}$ ligne doit donner le nombre minimal de différences que l'on peut trouver entre le code du $i^{\text{ème}}$ gène de l'entrée et un fragment de même longueur dans la grande chaîne d'ADN fournie.

Exemple

entrée :

```
↳ 3 6 12
   0 0 1 1 0 1
   1 1 1 1 1 1
   1 0 0 0 0 1
   1 0 1 1 0 0 0 0 1 1 0 0
```

sortie :

```
↳ 1
   3
   0
```

Commentaires

Premier motif : 1 différence au moins

```
1 0 1 1 0 0 0 0 1 1 0 0
      0 0 1 1 0 1
```

Second motif : 3 différences au moins

```
1 0 1 1 0 0 0 0 1 1 0 0
1 1 1 1 1 1
```

Troisième motif : 0 différence est possible

```
1 0 1 1 0 0 0 0 1 1 0 0
      1 0 0 0 0 1
```

[Aller résoudre ce problème](#)

Fragments d'ADN

