

TP5 : Programmation dynamique

Le but de ce TP est d'implanter le calcul de la distance d'édition vu en cours, ainsi que l'alignement de deux mots. Les fichiers suivants sont fournis :

- `matrice.h` et `matrice.cc` qui contiennent des fonctions pour déclarer, initialiser, désallouer et afficher une matrice (à ne pas modifier);
- `edition.h` (à ne pas modifier) et `edition.cc` (**seul fichier à modifier**) pour l'implantation des algorithmes vus en cours;
- le fichier de test `tests.cc` et le `Makefile` (à ne pas modifier);
- un dossier `covid19/` pour la question 6.

Le programme principal ne prend pas de paramètre (sauf à la question 6.) et permet de tester indépendamment chaque fonction. **Testez donc votre code à chaque question** (et pas uniquement avec les EXEMPLES encadrés)!

1. Compléter la fonction `int min(int a, int b, int c)` qui renvoie le minimum de ses trois arguments.

EXEMPLE

```

Entrer le numéro de question à tester (0 pour quitter) : 1
Entrer a : 2
Entrer b : 1
Entrer c : 6
min(2,1,6) = 1
  
```

2. Compléter la fonction `int valeur(int** E, int i, int j)` : cette fonction doit renvoyer `E[i][j]` si $i \geq 0$ et $j \geq 0$, $i+1$ si $j = -1$ et $j+1$ si $i = -1$.

EXEMPLE

```

Entrer le numéro de question à tester (0 pour quitter) : 2
Entrer le nb de lignes du tableau E : 5
Entrer le nb de colonnes du tableau E : 12
Entrer la ligne i : 3
Entrer la ligne j : 8
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
valeur(E, 3, 8) = 1
  
```

3. Compléter la fonction `int distanceEdition(string s1, string s2, int** E)` pour remplir la matrice E. On utilisera la fonction `valeur` pour éviter d'avoir à tester si $i - 1$ ou $j - 1$ est négatif. À ce stade, la fonction est censée calculer correctement la distance entre deux mots.

EXEMPLE

```

Entrer le numéro de question à tester (0 pour quitter) : 3
Entrer la première chaîne de caractère : AGORRYTNES
Entrer la seconde chaîne de caractère : ALGORITHMES
distanceEdition(AGORRYTNES, ALGORITHMES) = 5
  
```

4. Compléter la fonction `string alignement(string& s1, string& s2, int** E)` qui calcule un alignement des deux chaînes `s1` et `s2` : la fonction doit modifier `s1` et `s2` en insérant le caractère `_` (*underscore*) dans `s2` pour marquer une suppression, ou dans `s1` pour marquer une insertion. **Ne pas tenir compte dans cette question de la chaîne aligne qui est renvoyée.**

Note. La fonction `s.insert(i, ch)` insère la chaîne `ch` en position `i` de la chaîne `s` : par exemple si `s = "mer"` et qu'on applique `s.insert(1, "ang")`, `s` devient `"manger"`.

EXEMPLE

```
Entrer le numéro de question à tester (0 pour quitter) : 4
Entrer la première chaîne de caractère : AGORRYTNES
Entrer la seconde chaîne de caractère : ALGORITHMES
distanceEdition(AGORRYTNES, ALGORITHMES) = 5
alignement : A_GORRYTNES
             ALGORITHMES
```

5. Finir de compléter la fonction `alignement` pour construire la chaîne `aligne` : celle-ci doit être constituée des quatre caractères R (remplacement), I (insertion), S (suppression) et | (barre verticale, pour signifier que les deux lettres alignées sont égales), afin d'indiquer pour chaque caractère des deux chaînes alignées du type de transformation (ou absence de transformation) effectuée.

EXEMPLE

```
Entrer le numéro de question à tester (0 pour quitter) : 5
Entrer la première chaîne de caractère : AGORRYTNES
Entrer la seconde chaîne de caractère : ALGORITHMES
distanceEdition(AGORRYTNES, ALGORITHMES) = 5
alignement : A_GORRYTNES
             |I||RRRR||
             ALGORITHMES
```

6. On va enfin utiliser notre programme de manière utile ! Le dossier `covid19` contient les génomes complets de plusieurs coronavirus :

- `2019-nCoV_WH01.fa`, `2019-nCoV_WH03.fa` et `2019-nCoV_WH04.fa` sont des virus prélevés chez des patients chinois lors de l'épidémie de Covid-19 ;
- `bat-SL-CoVZC45.fa` et `bat-SL-CoVZXC21.fa` sont des virus portés par des chauve-souris ;
- `MERS-CoV.fa` et `SARS-CoV.fa` sont des virus prélevés chez des patients lors des épidémies de MERS (Moyen-Orient, 2012) et de SRAS (Asie, 2003).

Pour calculer la distance entre deux génomes contenus dans deux fichiers, il faut utiliser le programme principal de la manière suivante : `./tests <fichier1> <fichier2>`.

📎 À l'aide de ces fichiers, tenter de déterminer l'origine du virus actuel : mutation à partir d'un virus porté par les chauve-souris, ou mutation à partir d'un virus humain (MERS-CoV ou SARS-CoV).

Note. Vous pourrez comparer vos résultats avec l'étude dont sont tirés ces fichiers, disponible à l'adresse [https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736\(20\)30251-8.pdf](https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(20)30251-8.pdf).