

TP 2. Structures internes de données

le 30 mars 2017

Le travail sera effectué en binôme. Vous utiliserez le langage de votre choix (C, JAVA, C++, Python, etc....)¹.

On met à disposition un fichier `word2.txt`, dit de référence, avec plus de 200000 mots d'anglais utilisant uniquement des lettres minuscules, et des fichiers texte avec des mots (les mêmes que ceux que vous avez manipulés).

On essaie toujours de résoudre le même problème : détecter si quels mots d'un second fichier texte ne se trouvent pas dans le fichier de référence.

Dans ce nouveau TP on demande d'implémenter le calcul de cette différence d'ensembles par le biais d'un arbre préfixe (un trie) implémenté comme un enchainement de tableau de pointeurs.

Votre exécutable doit fonctionner sans avoir à saisir les données, il acceptera en entrée en ligne de commande un seul paramètre `fichier1.txt` et produira à la sortie standard les mots.

Voici les pas à mettre en place :

- définissez votre structure de données.
- initialisez votre trie qui ne contient rien.
- écrivez les fonctions d'insertion et de recherche.
- insérez un à un les mots du fichier de référence (pour la mise en place du programme prenez un tout petit fichier texte d'une dizaine de mots).

1. On conseille l'usage du langage C.

- faites rechercher un mot (un seul)
- faites que le programme final de recherche de tous les mots d'un fichier
- vérifiez que tout fonctionne bien pour le grand fichier de référence.

La date limite pour la remise du travail est le 3 avril 2018 au plus tard. La remise se fait par le biais de la plateforme Campus.

On dépose sur campus un fichier archive contenant son code source (ou ses codes, si plusieurs), un fichier de nom **README-eleve1-eleve2** qui décrit comment compiler et un autre fichier en format .pdf avec un (bref) compte rendu de TP. Il est fortement souhaitable que les programmes soient clairs, bien alignés et avec des commentaires. Dans l'archive ne mettez pas l'exécutable obtenu après la compilation, sauf si c'est un exécutable écrit en Shell, ne mettez pas non plus les fichiers texte avec les mots. Vous avez la possibilité de mettre les commentaires et de rédiger le compte rendu en français, en anglais ou en espagnol.

Si jamais vous avez une forte envie de continuer à programmer vous pouvez aussi essayer :

1. videz votre trie en affichant tous les mots qui s'y trouvent.
2. implémentez un Patricia trie