

TP 1. Structures internes de données

16 - 29 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 `word.txt`, dit de référence, avec plus de 200000 mots d'anglais, et des fichiers texte avec des mots (les mêmes que ceux que vous avez manipulés).

Il s'agit de détecter les mots issus d'un des fichiers texte qui ne se retrouve pas dans le fichier de référence.

Dans le TP on demande d'implémenter le calcul de cette différence d'ensembles par le biais d'une table de hachage. 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. Comptez aussi le nombre de collisions rencontrées et affichez le dans un petit fichier à part `coll.txt`.

Voici les pas à mettre en place :

- regardez bien les fichiers qui sont mis à votre disposition pour décider de la taille de la table de hachage, mettez dans le compte rendu de TP la justification (même sommaire) de vos choix
- implémentez une fonction de hachage qui prend en entrée une chaîne de caractères pas trop longue et qui fonctionne bien (bonnes propriétés générales). A la rigueur vous pouvez en implémenter deux.

1. On conseille l'usage du langage C.

- pensez à résoudre les collisions, justifiez votre choix et l'implémentation mise en place
- écrivez les fonctions d'insertion et de recherche
- dans le programme principal implémentez le calcul de la différence
- indiquez dans le compte rendu les choix faits, les caractéristiques de la table de hachage

La date limite pour la remise du travail est le 29 mars 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 d'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.

Afin de comparer les fonctions de hachage entre les divers auteurs une fonction de hachage prendra deux paramètres (un de type `char *` et un autre de type entier qui est la taille de la table de hachage) et retournera un entier. Si votre fonction de hachage demande d'autres paramètres ou moins ou plus, écrivez une qui appelle celle-ci avec une constante ou des constantes.

Un bonus sera accordée à la meilleure implémentation de la fonction de hachage.

Les programmes seront testés sur les fichiers fournis.