

Sujet : Indexation et recherche de mots contenus dans un ensemble de fichiers textes.

Méthode à suivre :

I
N
D
E
X
A
T
I
O
N

1. Construire un ensemble de mots dit vides. Les mots vides sont des mots que l'application doit ignorer, ce sont généralement les mots ayant une longueur égale à un et deux et certains de longueur trois tels que : « a à de la là le ci ce des les ». Cet ensemble représente les mots à ne pas indexer.
2. Pour chacun des fichiers du répertoire, réalisez une indexation dans un tableau de chaîne de caractères. Chaque chaîne de caractère du tableau doit contenir :

Mot	Fichier	Ligne	Position dans la ligne
-----	---------	-------	------------------------

Par mot on comprend tous les mots des fichiers sauf les mots vides.

3. Une fois tous les fichiers indexés, trie le tableau par ordre alphabétique sur les mots et enregistrez le tableau dans un fichier texte.

Re
ch
er
che

4. Dans la partie recherche il faut charger en mémoire le fichier index dans un tableau.
5. Recherchez la première occurrence du mot recherché par l'utilisateur dans l'index et afficher au minimum 4 résultats si le mot existe et plus de 4 résultats si l'utilisateur le souhaite. Le résultat consiste à afficher le nom du fichier contenant le mot, suivis de la ligne contenant le mot. Pour bien montrer le résultat le mot doit être affiché par exemple dans une autre couleur.

Question Bonus :

L'utilisateur peut rajouter un fichier dans le répertoire, le programme doit pouvoir l'indexer et rajouter les nouveaux mots à l'ancien index.

Permettre la recherche de plusieurs mots à la fois, et ou la tolérance à certaines erreurs de frappe

Points importants du TP

- Programmation en C sous code blocks, vous pouvez rajouter une bibliothèque graphique.
- La présentation doit être soignée (pas d'erreurs de français dans les menus et les commentaires).
- Les programmes doivent être bien commentés et bien organisés. (fonction, procédure, bibliothèque, indentation,...)
- Le TP doit être rendu en main propre sur cd, le répertoire du projet doit être la concaténation des noms du binôme suivi du numéro de groupe ex : bouali_benali_8.
- Travail en binôme obligatoire sauf si groupe impaire.