

L2 Informatique

TP Algorithmique – Recherche de mot dans un dictionnaire

Ce TP consiste à stocker l'ensemble de mots d'un dictionnaire¹ au moyen d'un arbre n-aire. Chaque descendant (fils) représente une lettre et chaque noeud un mot constitué des lettres du chemin menant de la racine vers le noeud en question. L'appartenance ou non d'un mot dans le dictionnaire est symbolisée par une valeur booléenne (étiquette du noeud). Le noeud racine représente le mot vide.

Exercice 1

On propose tout d'abord une structure d'arbres où chaque noeud possède 26 sous-arbres (certains éventuellement vides), un pour chaque lettre (de A à Z, pour simplifier on ne considère pas de caractères accentués). Il ne s'agit donc pas exactement d'un arbre général puisque c'est le numéro du descendant qui implique la lettre (A pour le premier fils, B pour le deuxième, *etc.*). Pour cela, on utilisera un tableau d'arbres pour représenter les 26 descendants d'un noeud. Les arêtes de l'arbre symboliseront donc des lettres ; ainsi, les caractères ne seront pas explicitement présents dans les noeuds.

Définir les types appropriés, et écrire un programme permettant les fonctionnalités suivantes :

1. Initialiser un dictionnaire (vide).
2. Ajouter un mot dans un dictionnaire.
3. Calculer la taille d'un dictionnaire (nombre de noeuds).
4. Ajouter tous les mots d'un dictionnaire figurant dans un fichier texte² (un mot par ligne).
5. Déterminer si un mot appartient à un dictionnaire.
6. Afficher les mots d'un dictionnaire correspondant à un *masque*. Un masque est un mot composé de lettres de l'alphabet et d'une ou plusieurs occurrences d'un caractère joker '?' pouvant remplacer n'importe quelle lettre de l'alphabet.

Exercice 2

Reprendre le programme en utilisant un arbre général implanté par une structure d'arbre binaire (représentation fils ainé / frère droit). Dans cette version, le k -ième descendant ne correspond pas nécessairement à la k -ième lettre de l'alphabet. Il faut donc que chaque noeud comporte le caractère de la lettre qu'il représente.

¹Dans ce TP, un *dictionnaire* ne fait pas référence à une structure de données tableau associatif, mais simplement à un ensemble de mots.

²Un fichier `larousse.txt` est disponible sur moodle.