

Arbre lexical fils gauche- frère droit

Dans toute la suite, on appelle **mot** une suite d'au plus 50 caractères ne contenant pas de séparateurs (chaîne au sens `scanf("%51s" . . .)`). Les accents sont donc significatifs.

("Interne" et "Interné" n'ont pas le même statut).

On appelle **lexique** d'un texte l'ensemble des mots qui apparaissent dans le texte.

1 Représentation du lexique par arbre binaire sous forme fils gauche,frère droit

On représentera le lexique à l'aide d'un arbre dans lequel chaque nœud possède 3 champs :

- un caractère
- 2 pointeurs `filsg`, `frered`

avec la structure C

```
typedef struct noeud{
unsigned char lettre;
struct noeud *filsg, *frered;
}Noeud, *Arbre;
```

Exemple :

l'arbre lexical de la figure 1 reconnaît l'ensemble de mots

{ce, ces, des, le, les, lettre, mes, mettre}.

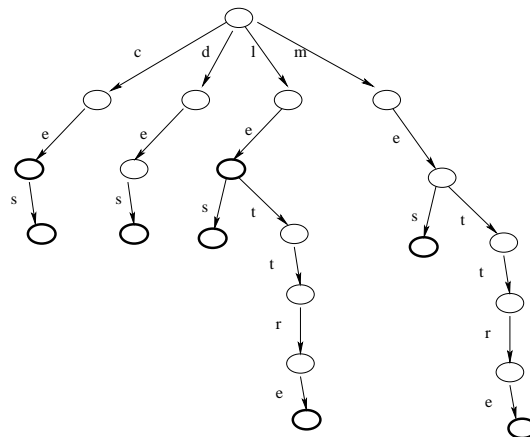


Figure 1: Arbre lexical

L'arbre sera construit sous la forme d'un arbre binaire fils gauche, frère droit, les nœuds frères étant ordonnés en ordre croissant voir figure 2. Un nœud pointe vers le premier de la liste de ses fils et vers son frère suivant.

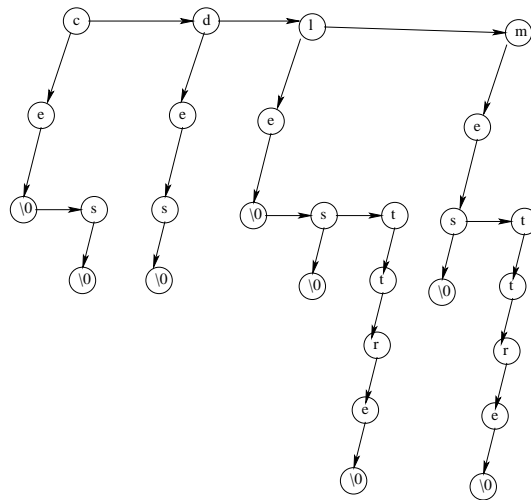


Figure 2: Arbre lexical forme fils gauche-frère droit

2 Fonctionnalités

Avec cette structure, construire le lexique d'un texte contenu dans un fichier fournit comme argument de la ligne de commande. L'exécutable aura pour nom **Lexique** et nécessitera toujours un argument : le nom du fichier à traiter.

Les fonctionnalités suivantes devront être implantées et accessibles par un menu:

1. afficher les mots du lexique en ordre alphabétique;
2. sauvegarder les mots du lexique en ordre alphabétique. Le nom du fichier de sauvegarde est le nom du fichier d'entrée suivi du suffixe(.L);
3. indiquer si le mot est présent dans le texte sous forme **present** ou **absent**;
4. sauvegarder l'arbre dans un fichier, le nom du fichier de sauvegarde est le nom du fichier d'entrée suivi du suffixe(.DICO)(format décrit plus loin);

Ces fonctionnalités seront également accessibles via une option de la ligne de commande (les options sont mutuellement exclusives):

1. **Lexique -l nom** affiche les mots du lexique du fichier **nom** en ordre alphabétique;
2. **Lexique -s nom** sauvegarde les mots du lexique du fichier **nom** en ordre alphabétique. Le nom du fichier de sauvegarde est le nom du fichier d'entrée suivi du suffixe(.L);
3. **Lexique -r Mot nom** indique si **Mot** apparaît dans le fichier **nom**

