

# Compression de données textes par la méthode de Huffman

## Félix Dallidet – Thibault Pichel

### Structure des modules:

hf_compress.c	hf_compress.h	types.h
hf_uncompress.c	hf_uncompress.h	

### Declaration des types :

Enregistrement freq\_table :  
    tableau d'entiers de 256 cases

Enregistrement Noeud :  
    frequence type entier  
    data type caractere  
    left type pointeur sur type Noeud  
    right type pointeur sur type Noeud

Enregistrement Dict :  
    codage type entier  
    data type caractere

### Raffinages hf\_compress:

#### **R0:**

Compresser un fichier par la methode de Huffman

#### **R1:**

- 1-Creer la table de frequence du fichier à compresser
- 2-Construire l'arbre de Huffman
- 3-Creer le dictionnaire
- 4-Generer le fichier compressé

#### **R2:**

##### **R2-1:**

Ouvrir le fichier  
Tant que tous les caracteres ne sont pas lus:  
    | Lire le caractere suivant  
    | Convertir le cractere en entier  
    | Incrementer de 1 la table de frequence a l'indice corespondant  
Fin tant que  
Fermer le fichier

##### **R2-2 :**

Creer un tableau dynamique de Noeuds pour chaque caractere present au moins 1 fois dans la table de frequence sans sous-arbres (descendants).  
Tant que le tableau compte plus d'un Noeud:  
    | Trier le tableau par « sum » croissante  
    | Remplacer les deux Noeuds les plus « à gauche » par un unique Noeud de «frequence» egale a la somme des «frequences» des 2 Noeuds precedants. Ce dernier pointe (par left) sur le « moins frequent » des 2 retirés, et par right sur l'autre.  
Fint tant que

##### **R2-3 :**

Dans l'arbre binaire de Huffman associer respectivement aux branches « left » et « right » les codes 0 et 1.  
Le code de chaque caractere est le nombre binaire obtenue en parcourant l'arbre depuis son sommet jusqu'au caractere.

Le dictionnaire est alors l'ensemble des codes ainsi obtenus.

**R2-3 :**

Créer un nouveau fichier

Ouvrir le nouveau fichier

Ecrire sur 4 octets la taille du texte décompressé

Ecrire sur 1 octet la taille de l'arbre

Ecrire l'ensemble des caractères différents utilisés dans le fichier à compresser dans leur ordre d'apparition dans l'arbre de Huffman (en partant d'en bas à gauche)

Ecrire le parcours infixe de l'arbre :

On part du sommet de l'arbre et on commence par indiquer le chemin du caractère le plus à gauche : on utilise des 0 pour préciser que l'on descend d'un Noeud, 1 pour monter d'un Noeud.

On parcourt ensuite l'arbre de gauche à droite en partant des positions Feuilles successives.

Ecrire le texte compressé

Raffinages hf\_uncompress:

**R0:**

Décompresser un fichier compressé par la méthode de Huffman

**R1:**

1-Retrouver l'arbre de Huffman grâce au parcours infixe

2-Retrouver le dictionnaire

3-Generer le fichier décompressé

**R2 :**

**R2-1 :**

Lire la taille du texte décompressé

Lire la taille de l'arbre sur le 5eme octet

Lire les caractères utilisés (du nombre de la taille de l'arbre)

Retrouver la longueur du parcours infixe (nbre de suites de 1)