

Algorithmique TD7

Magistère Informatique 1ère Année

22 Octobre 2008

Exercice 1 Adressage Ouvert

On prend une table de hachage où les collisions sont gérées par adressage ouvert. Le but est de calculer le nombre de sondages moyen effectué lors d'une recherche *fructueuse* ou *infructueuse*. On appelle m la taille de la table et n le nombre de clés qu'elle contient. Soit c une clé.

1. Exprimer le nombre de configurations pour lesquelles une recherche *infructueuse* de c effectue k sondages.
2. Montrer que :

$$\sum_{0 \leq i \leq l-1} C_i^{l-j-1} = C_i^j$$

En déduire le coût moyen d'une recherche *infructueuse* :

$$\frac{m+1}{m+1-n}$$

3. Exprimer le coût moyen d'une recherche *fructueuse* de c en fonction du coût des recherches infructueuses qui ont permis la construction de la table. En déduire que ce coût moyen est :

$$\frac{m+1}{n} (H_{m+1} - H_{m-n+1})$$

Où $H_j = 1 + \frac{1}{2} + \dots + \frac{1}{j}$, est le j ème nombre harmonique.

Exercice 2 Codage de Huffman, Algorithmes Gloutons

Soit Σ un alphabet fini et de cardinal supérieur ou égal à deux. On appelle mot de code toute suite finie de 0 et de 1, soit \mathcal{C} l'ensemble des mots de code. On appelle codage binaire une application injective α de l'alphabet Σ dans \mathcal{C} . En utilisant l'opération concaténation, α s'étend de manière naturelle en :

$$\alpha : \Sigma^* \rightarrow \mathcal{C}$$

On dit qu'un codage est de longueur fixe quand toutes les lettres sont codées par un mot de code de même longueur. Un codage est dit préfixe si aucune lettre n'est codée par un mot de code qui est préfixe du codage d'une autre lettre.

- 1) Montrer que pour un codage de longueur fixe et un codage préfixe, α est injective sur Σ^* .
- 2) Montrer qu'on peut représenter un codage préfixe par un arbre binaire dont les feuilles sont les lettres de l'alphabet.

On considère un mot w . A chaque codage préfixe de w , représenté par un arbre T , est associé un coût défini par :

$$C_w(T) = \sum_{a \in \Sigma} f(a)l_T(a)$$

avec $f(a)$ le nombre d'occurrences de a dans w et $l_T(a)$ la taille du mot de code associé à a par T . Un codage préfixe est dit optimal si il minimise la fonction C_w

3) Montrer qu'à un codage préfixe optimal correspond un arbre binaire où tout nœud interne a deux fils.

4) Montrer qu'il existe un codage préfixe optimal pour lequel les deux lettres dont le nombre d'occurrences est le plus faible sont sœurs dans l'arbre.

Etant donnés x et y les deux lettres dont le nombre d'occurrences est le plus faible dans w , on considère l'alphabet $\Sigma' = \Sigma - x, y + z$ où z est une nouvelle lettre à laquelle on associe $f(z) = f(x) + f(y)$.

5) Soit T' l'arbre d'un codage optimal pour Σ' , montrer que l'arbre T obtenu à partir de T' en remplaçant la feuille associée à z par un nœud interne ayant x et y comme feuilles représente un codage optimal pour Σ .

6) En déduire un algorithme recherchant un codage optimal et donner sa complexité.