Huffman

Quentin Januel, Loïc Mohin et Anthony Villeneuve

29 janvier 2019

Table des matières

1 Définitions

1.1 Alphabet

On appelle Σ un alphabet dont les éléments sont appelés des lettres.
Un mot sur Σ est un n-uplet de lettres : $m = (a_1,\ a_2,\ ...,\ a_n)$.
L'ensemble des mots sur Σ est noté $\Sigma^* := m \in \Sigma^n,\ \forall n \in \mathbb{N}$.
Soit $m = (a_1,\ a_2,\ ...,\ a_n)$ un mot, on appelle longueur du mot m notée $|m|$ l'entier n.
Enfin, on note ε le mot vide (unique mot de longueur 0).

On peut munir Σ^* d'une loi de composition interne, la concaténation $+$:
$(a_1,\ ...,\ a_n) + (b_1,\ ...,\ a_n) = (a_1,\ ...,\ a_n,\ b_1,\ ...,\ a_n)$.
On observe alors que $(\Sigma^*, +)$ est un monoïde.

1.2 Arbre binaire

Soit E un ensemble, on appelle alors a un arbre binaire sur E un triplet (G, x, D) avec x un élement de E et G, D des arbres binaires.
On le définit par récurrence ainsi :
- L'arbre vide, noté a_0, est un arbre binaire,
- Pour tout G, D arbres binaires, pour tout $x \in E$, $(G,\ x,\ D)$ est un arbre binaire.

On dit que G (resp. D) est le sous arbre gauche (resp. droit) de $a = (G,\ x,\ D)$ noté $g(a)$ (resp. $d(a)$). x est le noeud de a, noté $n(a)$.
On note aussi $\mathcal{A}_E$ l'ensemble des arbres binaires sur E.

1.3 Arbre de Huffman

On appelle arbre de Huffman un arbre binaire a sur $\mathbb{N}$ qui respecte les propriétés suivantes :
- $n(a) = n(g(a)) + n(d(a))$,
- $g(a)$ et $d(a)$ sont tous deux des arbres de Huffman,
- ... (il manque un truc ici, à faire !).

On note $\mathcal{H}$ l'ensemble des arbres de Huffman et on a $\mathcal{H} \subset \mathcal{A}_{\mathbb{N}}$.

2 Classes d'équivalence d'arbres de Huffman

Intuitivement, on aimerait définir des classes d'équivalences des arbres de Huffman selon la longueur de la chaine de bits qu'ils encodent. Tâchons donc d'abord de pouvoir évaluer cette valeur à l'aide d'une fonction :

$$\begin{array}{rccc} \omega : & \mathcal{H} & \to & \mathbb{N} \\ & a & \mapsto & \begin{cases} 0 \text{ si } a = a_0 \\ n(a) + \omega(g(a)) + \omega(d(a)) \text{ sinon} \end{cases} \end{array}$$

Nous pouvons à présent définir $\mathcal{R}$ notre relation d'équivalence. Soient $a, b \in \mathcal{H}$, on a alors

$$a\mathcal{R}b \Longleftrightarrow \omega(a) = \omega(b)$$

On dénote également $\overline{\mathcal{H}} := \mathcal{H}/\mathcal{R}$ l'ensemble quotient de $\mathcal{H}$ par $\mathcal{R}$.