TD 7 : Fonctions de hachage

Exercice 1. *Adressage ouvert*

Plutôt que résoudre les conflits par chaînage, c'est-à-dire faire en sorte que chaque case de la table ne contienne pas uniquement un élément mais un ensemble d'éléments, l'*adressage ouvert* consiste à résoudre les conflits en allant chercher une autre case de la table. Formellement, on suppose disposer de m fonctions de hachages h_0, …, h_{m-1} de U dans $\{0,\dots,m-1\}$. Pour insérer un élément x, on l'insère en case $T[h_0(x)]$ si elle est libre, sinon en case $T[h_1(x)]$ si elle est libre, et ainsi de suite. Si aucune des cases $T[h_i(x)]$ n'est libre, on déclare la table remplie et on renvoie une erreur.

On se place dans le modèle aléatoire des fonctions de hachages, avec une hypothèse en plus [1] : on suppose que pour tout x, $(h_0(x),\dots,h_{m-1}(x))$ est une permutation aléatoire de $\{0,\dots,m-1\}$. On souhaite borner l'espérance $\mathbb{E}_{m,n}$ du nombre d'essais pour insérer un nouvel élément x dans une table de taille m contenant déjà n éléments.

1. Montrer que pour tout nouvel élément x, la probabilité que $T[h_0(x)]$ soit libre est $1-n/m$.
2. Montrer que $\mathbb{E}_{m,n} = 1 + \frac{n}{m}\mathbb{E}_{m-1,n-1}$.
3. En déduire que $\mathbb{E}_{m,n} \leq m/(m-n)$.

Exercice 2. *Une famille quasi-universelle*

On s'intéresse à une fonction de hachage très efficace à calculer en pratique. Pour tout entier impair a, on pose

$$h_a(x) = \left\lfloor \frac{(ax)\%2^w}{2^{w-\ell}} \right\rfloor$$

où w et ℓ sont des paramètres que l'on suppose fixés et $x\%y$ représente le reste dans la division euclidienne de x par y. En particulier, $x\%2^w$ représente l'entier obtenu en ne gardant que les w bits de poids faible de x.

1. Montrer que pour tout a et tout x, $h_a(x)$ est un entier compris entre 0 et $2^\ell - 1$.
2. On voit chaque entier comme un tableau de bits.
   - **i.** Quelle est la taille de ax, si a et x sont de taille w ?
   - **ii.** Exprimer le tableau représentant $h_a(x)$ comme un sous-tableau de ax.
   - **iii.** En supposant que w est la taille des entiers sur votre ordinateur, donner une implantation efficace de h_a qui prend a, x et ℓ en paramètres. *Utiliser par exemple le langage C.*

On note W l'ensemble des entiers impairs entre 0 et $2^w - 1$.

3. Montrer que pour tout couple $(x,y) \in W^2$, il existe un unique $a \in W$ tel que $(ax)\%2^w = y$.
4. On va montrer que la famille $(h_a)_{a\in W}$ est presque-universelle.
   - **i.** On fixe $a \in W$. Montrer que $h_a(x) = h_a(y)$ si et seulement si $h_a(x-y) = 0$ ou $h_a(y-x) = 0$.

   On écrit $(y-x)\%2^w = q2^r$ où q est impair.

   - **ii.** Montrer que les $r+1$ bits de poids faible de $(aq2^r)\%2^w$ sont $10\cdots0$.
   - **iii.** Montrer que si $r \geq w-\ell$, $\Pr[h_a(y-x)=0] = 0$.
   - **iv.** Montrer que si $r < w-\ell$, $\Pr[h_a(y-x)=0] = 1/2^\ell$.
   - **v.** Conclure.

1. Appelée *hypothèse forte du hachage uniforme*.