HW3 BAsic Algorithms

Algo L3

6.10.2016 - **Due on Thursday 13.10 before 8:30**

Please send me your solutions as a PDF file named « HW3-BOTH_YOUR_NAMES.pdf » at:
Cours.AlgoL3@ens.fr
with Subject « [HW3] »
(or return it at the next lecture) on **Thursday 13.10 before 8:30**.

■ **Exercise 1 (A streaming algorithm for the second moment of the frequencies).** We are given a stream of numbers $x_1, \dots, x_n \in \{0, \dots, m-1\}$ and we want to compute the sum of the squares of the frequencies of each values 0 to $m-1$ in this stream: if $f_a(x) = \#\{i : x_i = a\}$, we want to compute $F_2(x) = \sum_{a=0}^{m-1} f_a(x)^2$.

Take a random function $h : \{1, \dots, m\} \to \{-1, 1\}$, i.e.: for all a, $h(a)$ is chosen at independently and uniformly at random in $\{-1, 1\}$. And do the following:

Algorithm 1 Second frequency moment random algorithm

Pick a hash function $h : \{0, \dots, m-1\} \to \{-1, 1\}$ uniformly at random
Compute $Z = h(x_1) + \cdots + h(x_n)$ while reading the stream
Output Z^2

► ***Question* 1.1)** *Show that $\mathbb{E}[Z^2] = F_2(x)$ where the expectation is taken over all the possible values for h.* ▷ Hint. *For $a \neq b$, show that $\mathbb{E}_h[h(a)h(b)] = 0$ and $\mathbb{E}_h[h(a)^2] = 1$.*

► ***Question* 1.2)** *Show that* $\mathbb{V}ar(Z^2) = \mathbb{E}[Z^4] - \mathbb{E}[Z^2]^2 = 2\sum_{a \neq b} f_a(x)^2 f_b(x)^2 \leqslant 2F_2(x)^2$.

Remark that this algorithm requires a lot of memory to store h: $O(m \log m)$ bits, almost as much as counting the frequencies independently ($O(m \log n)$ bits). But remark that we only need the values of h to be 4-*wise* independent to obtain the results above. Let us thus use the following construction for h that will require much less memory.

Consider the field $\mathbb{F}_{2^k}$ where $k = \lceil \log_2 m \rceil$ such that $2^{k-1} < m \leqslant 2^k$. Let us identify the elements of $\mathbb{F}_{2^k}$ as a string of k bits and as numbers from 0 to $2^k - 1$ as well. Let $\pi : \mathbb{F}_{2^k} \to \{-1, 1\}$ be the function that associates to any number $a \in \mathbb{F}_{2^k}$ the value -1 if the first bit of a is 0 and the value $+1$ otherwise.

For all 4-tuple $(u, v, w, t) \in (\mathbb{F}_{2^k})^4$, let $P_{uvwt} : \mathbb{F}_{2^k} \to \mathbb{F}_{2^k}$ be the polynomial:

$$P_{uvwt}(a) = ua^3 + va^2 + wa + t,$$

and set $h_{uvwt}(a) = \pi(P_{uvwt}(a))$.

► ***Question* 1.3)** *Show that if u, v, w, t are chosen independently and uniformly at random in $\mathbb{F}_{2^k}$, then for all fixed* distinct *values $a, b, c, d \in \mathbb{F}_{2^k}$, the random 4-tuple $(P_{uvwt}(a), P_{uvwt}(b), P_{uvwt}(c), P_{uvwt}(d))$ is uniform in $(\mathbb{F}_{2^k})^4$.*
▷ Hint. *Recall van der Mond matrices...*

► ***Question* 1.4)** *Conclude that when u, v, w, t are chosen independently and uniformly at random in $\mathbb{F}_{2^k}$, the values $h_{uvwt}(0), \ldots, h_{uvwt}(m-1)$ are 4-wise independent uniform random variables with values in $\{-1, 1\}$.*