Magistère Informatique et Modélisation
(Ens-Lyon, Université Lyon-I)

Calcul formel et cryptographie

Contrôle du vendredi 14 mai 2004 . Corrigé

Exercice 1 : Question de cours

1. Sur un anneau de caractéristique 2 l'application $t \mapsto t^2$ est linéaire. D'ou $X^4+1 = (X^2)^2+1^2 = (X^2+1)^2 = X^4+1$.
2. La dérivée de F, $4X^3$, n'a pas de facteur commun avec F qui est donc sans facteur multiple. La matrice de Berlekamp est la matrice des coefficients des vecteurs $\overline{1}, \overline{X^p}, \overline{X^{2p}}, \overline{X^{3p}}$, dans la base $\overline{1}, \overline{X}, \overline{X}^2, \overline{X}^3$, de l'anneau quotient $\mathbb{F}_2[X]/(F)$. L'équation $X^4+1 \equiv 0 \pmod F$ donne $\overline{X}^4 = -1$ puis $\overline{X}^8 = 1$. Comme $p-1$ est un multiple de 8, on en déduit $\overline{X}^{p-1} = 1$ et enfin $\overline{X}^p = \overline{X}$. La matrice de Berlekamp de F est la matrice identité. Le sous-espace propre relatif à la valeur propre 1 est donc de dimension 4, et F est le produit de 4 facteurs irréductibles.
3. Si p est congru à 3 modulo 8, l'équation $\overline{X}^8 = 1$ donne $\overline{X}^{p-3} = 1$, et $\overline{X}^p = \overline{X}^3$. Puis $\overline{X}^{2p} = \overline{X}^6 = -\overline{X}^2$ et enfin $\overline{X}^{3p} = -\overline{X}^2\overline{X}^3 = \overline{X}$. La matrice de Berlekamp est donc

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \text{ et } M - \mathrm{I}_d = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix},$$

est de rang 2 (la dernière ligne est l'opposée de la deuxième). Son noyau est de dimension 2. Le polynôme F a donc deux facteurs irréductibles.

Exercice 2 : Un autre regard sur le paradoxe des anniversaires

L'espace probabilisé Ω est l'ensemble produit $[N]^n$ en notant $[N] = \{1, 2, \cdots, N\}$, muni de la probabilité uniforme. Le cardinal de Ω est N^n, la probabilité de chaque singleton $\{(x_1, x_2, \cdots, x_n)\}$ de $\mathcal{P}(\Omega)$ est donc N^{-n}.

1. X est la somme des $X_{i,j}$.
2. (a) L'espérance étant linéaire on en déduit $E(X) = \sum_{1\leq i<j\leq n} E(X_{i,j})$ L'évènement $X_{i,j} = 1$ est l'ensemble des n-uplets $(x_1, x_2, \cdots, x_n)$ tels que $x_i = x_j$. Se donner un tel n-uplet c'est se donner arbitrairement N^{n-1} valeurs dans $[N]$. La probabilité de l'évènement $X_{i,j} = 1$ est donc $p = N^{n-1}N^{-n} = 1/N$. La variable $X_{i,j}$ ne prenant que la valeur 1 son espérance est p. On en déduit

$$E(X) = \sum_{1\leq i<j\leq n} p = \frac{n(n-1)}{2}p = \frac{n(n-1)}{2N}.$$

   (b) Pour $n = N$ cela donne $E(X) = \frac{1}{2}(N-1)$, et pour $N = 365$, $n = 40$ cela donne $E(X) = \frac{40\times 39}{365} \approx 2,137$.

   (c) Pour que $E(X) \geq 1$, il faut et il suffit que $n(n-1) \geq 2N$, soit $n - \frac{1}{2} \geq \sqrt{2N + \frac{1}{4}}$.
3. (a) Si $N \geq 3$ les variables $X_{i,j}$ ne sont pas indépendantes car, par exemple, l'évènement $(X_{1,2} = 1) \cap (X_{2,3} = 1) \cap (X_{1,3} = 0)$, est de probabilité nulle, et donc distincte du produit des probabilités $P(X_{1,2} = 1)P(X_{2,3} = 1)P(X_{1,3} = 0) = 1/N^3$.

   (b) Les variables $X_{i,j}$ sont néanmoins deux à deux indépendantes. Cela est évident dans le cas de $X_{i,j}$ et $X_{i',j'}$ avec i, j, i', j' tous distincts, car les résultats des tirages de rang i, j, i', j' sont indépédants (parce qu'il s'agit de tirages avec remise). Il reste à démontrer que, par exemple, sans nuire à la généralité, les 2 variables $X_{1,3}$ et $X_{2,3}$ sont indépendantes. Il faut démontrer que, pour chaque couple $(\varepsilon_1, \varepsilon_2)$ de l'ensemble $\{0, 1\}^2$, on a

$$P((X_{1,3} = \varepsilon_1) \cap (X_{2,3} = \varepsilon_2)) = P(X_{1,3} = \varepsilon_1)P(X_{2,3} = \varepsilon_2).$$

      i. Si $(\varepsilon_1, \varepsilon_2) = (1, 1)$, l'évènement $X_{1,3} = X_{2,3} = 1$ est constitué des n-uplets $(x_1, x_2, x_3, \ldots, x_n)$ tels que $x_1 = x_2 = x_3$. Le nombre de ces n-uplets est N^{n-2} car un tel n-uplet est caractérisé par la donnée, arbitraire, des valeurs de $x_3, x_4, \cdots, x_n$. On a donc

$$P(X_{1,3} = 1) \cap (X_{2,3} = 1)) = N^{n-2}N^{-n} = 1/N^2 = P(X_{1,3} = 1)P(X_{2,3} = 1).$$

      ii. Si $(\varepsilon_1, \varepsilon_2) = (0, 0)$, l'évènement $X_{1,3} = X_{2,3} = 0$ est constitué des n-uplets $(x_1, x_2, x_3, \ldots, x_n)$ tels que $x_1 \neq x_3$ et $x_2 \neq x_3$. Pour obtenir un tel n-uplet on se donne arbitrairement x_3, puis x_1 et x_2 arbitraires dans $[N] \setminus \{x_3\}$,

et enfin $x_4, \cdots, x_n$ arbitraires dans $[N]$. Le cardinal de l'évènement $X_{1,3} = X_{2,3} = 0$ est donc $(N-1)^2 N^{n-2}$, et sa probabilité $(N-1)^2 N^{n-2} N^{-n} = (1-\frac{1}{N})^2$. On a donc

$$P((X_{1,3}=0) \cap (X_{1,3}=0)) = \left(1-\frac{1}{N}\right)^2 = P(X_{1,3}=0)P(X_{2,3}=0).$$

iii. Si $(\varepsilon_1, \varepsilon_2) = (1,0)$, l'évènement $(X_{1,3}=1)$ et $(X_{2,3}=0)$ est constitué des n-uplets $(x_1, x_2, x_3, \ldots, x_n)$ tels que $x_1 = x_3$ et $x_2 \neq x_3$. Pour obtenir un tel n-uplet on se donne arbitrairement $x_1 = x_3$, ce qui laisse $N-1$ choix pour x_2, et enfin N choix pour chacun des $x_j, j \geq 4$. Cet évènement est donc de cardinal $N(N-1)N^{n-3}$ et sa probabilité est $(N-1)N^{-2} = (1-1/N)1/N$. On a donc

$$P((X_{1,3}=1) \cap (X_{2,3}=0)) = \frac{1}{N}\left(1-\frac{1}{N}\right) = P(X_{1,3}=1)P(X_{2,3}=0).$$

iv. Le dernier cas $(\varepsilon_1, \varepsilon_2) = (0,1)$ se traite comme le précédent.

(c) Les $X_{i,j}$ sont deux à deux indépendantes. La variance de leur somme est donc la somme de leurs variances. Comme $X_{i,j}$ ne prend que les valeurs 0 et 1, on a $X_{i,j}^2 = X_{i,j}$ et

$$V(X_{i,j}) = E(X_{i,j}^2) - E(X_{i,j})^2 = E(X_{i,j}) - E(X_{i,j})^2 = \frac{1}{N} - \frac{1}{N^2}.$$

Il en résulte

$$V(X) = \frac{n(n-1)}{2}\frac{1}{N}\left(1-\frac{1}{N}\right) = \left(1-\frac{1}{N}\right)E(X).$$

4. Si $X = 0$, alors $|X - E(X)| = E(X) \geq E(X)$. L'évènement $X = 0$ est donc contenu dans l'évènement $|X - E(X)| \geq E(X)$ et, par l'inégalité de Tchebichev, en y remplaçant λ par $E(X)$, sa probabilité est majorée par

$$P(X=0) \leq \frac{V(X)}{E(X)^2} = \left(1-\frac{1}{N}\right)E(X)/E(X)^2 = \left(1-\frac{1}{N}\right)/E(X) = \frac{2(N-1)}{n(n-1)}.$$

Pour que cette probabilité soit majorée par $1/2$ il suffit que $4(N-1) \leq n(n-1)$.

5. **Une application du paradoxe des anniversaires.** Pour obtenir une collision, on tire au hasard des éléments x dans E, en calculant à chaque fois $f(x)$, et en **rangeant dans un dictionnaire** la valeur obtenue si elle n'y figure pas déja, l'information associée à l'entrée $f(x)$ du dictionnaire est l'antécédent x de $f(x)$. Par le paradoxe

des anniversaires le nombre de valeurs à calculer avant d'obtenir une collision est de l'ordre de $\sqrt{\text{card}(K)} \approx 2^{32} \approx 4.3 \times 10^9$. Ce nombre est donc le nombre attendu de valeurs à calculer, et aussi le nombre de valeurs à stocker dans le dictionnaire, c'est à dire la place mémoire nécessaire. **Attention** Il ne suffit pas de ranger dans une liste ou un tableau, les valeurs obtenues, car, après l'obtention de la k^{eme} valeur, $f(x_k)$, il faudrait parcourir les $k-1$ premières cases du tableau pour s'assurer que la valeur $f(x_k)$ n' a pas encore été obtenue. Si n est la première valeur de k pour laquelle on obtient une valeur déja obtenue, le nombre de comparaisons effectuées serait $1+2+3+\cdots+(n-1) = n(n-1)/2$. Comme n est de l'ordre de 2^{32} ce nombre est de l'ordre de 2^{63} ce qui est de l'ordre de 10^{20} au lieu de 10^{10}.

Problème

Partie I

1. Soit m un multiple de $p-1$ et de $q-1$, et x premier avec n. Alors x est premier avec p, et, par le petit théorème de Fermat $x^m \equiv 1 \pmod{p}$. De même $x^m \equiv 1 \pmod{q}$. x^m est donc une solution du système

$$\begin{cases} u & \equiv & 1 & (\text{mod } p) \\ u & \equiv & 1 & (\text{mod } q) \end{cases}$$

   Par le théorème des restes chinois ce système a une unique solution modulo $n = pq$. Une solution évidente est $u = 1$. On a donc $x^m \equiv 1 \pmod{n}$, et m est un bon exposant.

   Réciproquement, soit m un bon exposant, et a (resp. b) un générateur du groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^\times$ (resp. $(\mathbb{Z}/q\mathbb{Z})^\times$), et, par le théorème des restes chinois, x un entier congru à a modulo p, et congru à b modulo q. La congruence $x^m \equiv 1 \pmod{n}$ implique la congruence $x^m \equiv 1 \pmod{p}$ soit $a^m \equiv 1 \pmod{p}$. Comme a est un générateur modulo p, m est un multiple de $p-1$. De même c'est un multiple de $q-1$.

2. L'ensemble des entiers x tels que $x^m \equiv 1 \pmod{n}$ est un sous-groupe de $(\mathbb{Z}/n\mathbb{Z})^\times$. Si m n'est pas un bon exposant ce groupe est un sous-groupe propre. Son cardinal est donc un diviseur strict de $\varphi(n)$, son cardinal est donc majoré par $\frac{1}{2}\varphi(n)$. Le cardinal de son complémentaire est donc au moins $\frac{1}{2}\varphi(n)$.

3. L'algorithme évident est le suivant

```
booleen BonExposant(entiers m,N,k)
Pour i de 1 a k faire
   Choisir x au hasard
   Si x^m ≢ 1 (mod n) renvoyer FAUX
  finpour ;
renvoyer VRAI
fin
```

La question, telle qu'elle était posée était incorrecte, puisqu'elle demandait de prouver un résultat faux. Ce qui est vrai, c'est que : *Sachant que m n'est pas un bon exposant, alors la probabilité de l'évènement " BonExposant renvoie VRAI " est majorée par* $1/2^k$. En effet sachant que m n'est pas un bon exposant, chaque fois que l'on tire x la probabilité que $x^m \equiv 1 \pmod{n}$ est au plus $1/2$. La probabilité que k répétions de cette expérience conduisent au ême résultat est donc majorée par $1/2^k$.

Partie II

1. Puisque $p-1$ est pair (ou $q-1$), et m multiple de $p-1$, il est pair. Fixons par exemple $k = 20$, et lançons l'appel `BonExposant(m/2,N,20)`. Si cet appel renvoie FAUX m est un bon exposant, est $m/2$ n'en est pas un. Si cet appel renvoie vrai, $m/2$ il est très probable que $m/2$ soit encore un bon exposant, on remplace m par $m/2$ et on recommence. Rapidement on finira par obtenir un entier m qui est très probablement un bon exposant, tandis que $m/2$ n'est plus un bon exposant.
   – Supposons $m/2$ multiple de $p-1$ et non multiple de $q-1$. Par le théorème des restes chinois l'application

   $$\overline{x} \mapsto (x_1 = x \mod p, x_2 = x \mod q)$$

   est une bijection de $(\mathbb{Z}/N\mathbb{Z})^\times$ sur $(\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times$. De plus, pour tout x_1 on a $x_1^{m/2} \equiv 1 \pmod{p}$ car $m/2$ est un multiple de $p-1$, et x_1 premier avec p. Tandis que, par le critère d'Euler, $x_2^{m/2} \equiv 1 \pmod{q}$ si, et seulement si x_2 est un carré modulo q c'est à dire pour exactement la moitié des valeurs de x_2.
   – Si $m/2$ n'est divisible ni par $p-1$, ni par $q-1$. Quelle que soit le couple $(\varepsilon_1, \varepsilon_2) \in \{-1, 1\}^2$, la moitié des valeurs de x_1 vérifient $x_1^{m/2} \equiv \varepsilon_1 \pmod{p}$, et la moitié des valeurs de x_2 vérifient $x_2^{m/2} \equiv \varepsilon_2 \pmod{q}$. Ainsi exactement le quart des entiers x dans $(\mathbb{Z}/N, \mathbb{Z})^\times$ sont tels que

   $$x^{m/2} \equiv \varepsilon_1 \pmod{p}, \qquad x^{m/2} \equiv \varepsilon_2 \pmod{q}.$$

2. Supposons donc que m est un bon exposant mais pas $m/2$. D'après la question précédente, pour la moitié des valeurs de x on a $x^{m/2} \mod p \not\equiv x^{m/2} \mod q$. En tirant au hasard quelques valeurs de x, on obtiendra rapidement un x tel que, par exemple,

$$x^{m/2} \equiv 1 \pmod{p} \qquad \text{et} \qquad x^{m/2} \equiv -1 \pmod{q},$$

Il en résulte que $x^{m/2} - 1$ est divisible par p, mais pas par q. Ainsi $x^{m/2} - 1$ est divisible par l'un des nombres p et q, mais pas par les deux. Il suffit de calculer le pgcd de $x^{m/2} - 1$ et de N pour obtenir, soit le facteur p, soit le facteur q de N.

```
procedure Factorise(N,m)
tant que BonExposant(m,n,20) faire m := m/2 fintanque
repeter
   Choisir x au hasard
tant que x^{m/2} mod p = x^{m/2} mod q.
renvoyer pgcd(x^{m/2} - 1, N)
fin
```

Cet algorithme probabiliste échouera si, lors de la phase initiale, une valeur $m/2$ qui n'était pas un bon exposant a été déclarée à tort bon exposant par la procédure `BonExposant`. Dans ce cas elle renverrait 1 au lieu de renvoyer un facteur de N. On a vu que cet évènement est très improbable. Il suffirait alors de recommencer. Elle pourrait aussi échouer, si dans la deuxième partie, qui est un jeu de pile ou face, on ne voyait jamais apparaitre pile !

Partie III

1. d est un exposant de déchiffrement pour la clef publique (N, e), alors $ed - 1$ est un bon exposant modulo N. Connaissant N, e, d on applique la procedure Factorise avec les arguments N et $ed - 1$ pour factoriser N.
2. Cela montre qu'une solution du problème du calcul d'un exposant de déchiffrement résoud le problème de la factorisation de N. C'est donc un problème aussi difficile que la factorisation.
3. Mais cela ne prouve pas que le problème du décryptage de RSA soit aussi difficile que la factorisation, car on ne sait pas prouver qu'un algorithme de décryptage donnerait un algorithme de calcul d'un exposant de déchiffrement. Contrairement à la situation du chiffrement de Rabin, qui est donc plus solide que RSA.