CRYPTOGAPHIE

1 Objectifs généraux

L'objectif de la cryptographie est de permettre la circulation sous forme cachée ou chiffrée C d'un message en clair M entre deux personnes, traditionnellement appelées Alice et Bob de telle sorte qu'une troisième personne, appelée Oscar, non autorisée ne puisse pas retrouver M à partir de C dans un temps "raisonnable"

La transformation de M en C appelée **chiffrement** se fait par l'intermédiaire d'une fonction (au sens mathématique et algorithmique aussi) de chiffrement E (encryption) telle que $C = E(M)$

Le **déchiffrement** de C en M se fait par l'intermédiaire d'une fonction D telle que $M = D(C)$

Par conséquent $D(E(M)) = M$

Définition 1 *$f : X \to Y$ est injective lorsque $\forall x, y \in X$ si $x \neq y$ alors $f(x) \neq f(y)$*
(ou par contraposition lorsque $\forall x, y \in X$ si $f(x) = f(y)$ alors $x = y$)
$f : X \to Y$ est surjective lorsque $\forall y \in Y \ \exists x \in X$ tel que $f(x) = y$

Exemples

1. La fonction $\mathbb{R} \ni x \to x^2 \in \mathbb{R}$ n'est pas injective n'est pas surjective, car 4 a deux antécédents -2 et 2, et les nombres négatifs n'ont pas d'antécédents
2. La fonction $\mathbb{R}^+ \ni x \to x^2 \in \mathbb{R}^+$ est injective et surjective

Théorème 1 *Si $f : X \to Y$ et $g : Y \to X$ sont telles que $\forall x \in X \ g(f(x)) = x$*
alors f est injective et g est surjective
Par conséquent E est injective et D surjective

Preuve

1. Soient x_1 et x_2 appartenant à X
   Si $f(x_1) = f(x_2)$ alors $g(f(x_1)) = g(f(x_2))$ or $g(f(x_1)) = x_1$ et $g(f(x_2)) = x_2$ donc $x_1 = x_2$ donc f est injective
2. Soit x un élément quelconque de X, montrons qu'il existe $y \in Y$ tel que $g(y) = x$
   En effet $y = f(x)$ car $g(y) = g(f(x)) = x$ donc $g : Y \to X$ est surjective

Définition 2 *$f : X \to Y$ est bijective si elle est injective et surjective*
Dans ce cas on peut définir la fonction réciproque de f, notée f^{-1} par $f^{-1} : Y \ni y \to x \in X$ tel que $f(x) = y$

Exemple

la fonction **exponentielle** est une bijection de $\mathbb{R}$ dans $\mathbb{R}^+$ et sa fonction réciproque est la fonction logarithme népérien

Théorème 2 *$f \circ f^{-1} = I_Y$ et $f^{-1} \circ f = I_X$ où $I_X : X \ni x \to x \in X$*

2 Système cryptographique à clé secrète

Les relations fondamentales s'écrivent avec un paramètre K appelé **clé**
$C = E_K(M)$ et $M = D_K(C)$
Cette clé est **partagée secrètement entre Alice et Bob**
Alice et Bob conviennent aussi d'un algorithme de chiffrement et de déchiffrement
La "sécurité de ce système" repose sur la clé qui est acheminée entre Alice et Bob par un "canal sécurisé"

Exemple 1 : Chiffrement affine

Les messages sont chiffrés à partir de l'alphabet latin (26 caractères non accentués) vers ce même alphabet
Les autres symboles, (par exemple virgule, apostrophe etc...sont laissés tel quel dans le message)

A chaque caractère est associé un nombre de 0 à 25, les restes de la division par 26 dans $\mathbb{Z}$
Par la suite on confond un caractère et ce nombre

On note $\mathbb{Z}/26\mathbb{Z}$ cet ensemble
On calcule dans cet ensemble modulo 26, et voici le principe du chiffrement affine :

On note $\mathbb{Z}/26\mathbb{Z}^*$, **l'ensemble des inversibles** de $\mathbb{Z}/26\mathbb{Z}$
La fonction de chiffrement d'un caractère à partir de la clé $K = (a, b)$ où $a \in \mathbb{Z}/26\mathbb{Z}^*$ et $b \in \mathbb{Z}/26\mathbb{Z}$, est définie par

$E_{(a,b)} \ : \ \mathbb{Z}/26\mathbb{Z} \ni x \to ax + b \in \mathbb{Z}/26\mathbb{Z}$

Le chiffrement d'un message revient à concaténer les chiffrements des caractères constituant le message :

Par exemple pour chiffrer le message en clair $M = bob$ on considère la suite 2,14,2 puis $E_{(a,b)}(2), E_{(a,b)}(14), E_{(a,b)}(2)$ que l'on traduit ensuite en caractères

1. Expliciter $\mathbb{Z}/26\mathbb{Z}^*$
2. Justifier que $K = (15, 7)$ est une clé possible
3. Chiffrer le message "bob" avec cette clé
4. Expliciter la fonction de déchiffrement $D_{(15,7)}$
5. Déchiffrer "jrlhc" avec la clé (15,7)

(Voir la suite en exos)

Exemple 2 : Chiffrement de Hill (voir exos)

3 Système cryptographique à clé publique

"*Comme nous, Ralph était un peu fou. Il faut être fou pour mener jusqu'au bout une recherche originale, et seuls des fous la poursuivront envers et contre tout. Vous avez une première idée, vous vous enthousiasmez, et ça ne marche pas. Alors vous passez à l'idée numéro 2, vous vous enthousiasmez, et ça ne marche pas. Alors vous avez l'idée 99, vous vous enthousiasmez, et ça ne marche pas. Seul un fou sera encore enthousiaste à sa centième idée, et il faudra peut-être avoir essayé cent voies avant que l'une conduise quelque part. A moins d'être assez fou pour retrouver à chaque fois tout votre enthousiasme, vous vous découragerez, et n'aurez pas l'énergie de mener les*

choses jusqu'au bout. Dieu bénit les fous."(Martin Hellman, professeur d"informatique à l'université de Stanford - Californie (1975))

3.1 Principe de Kerckhoffs (1883)

Jusqu'à présent pour échanger des messages cryptés l'émetteur et le récepteur devaient partager un secret.

C'est le principe de Kerckhoffs(1883) :

"La sécurité d'un système de chiffrement ne doit pas dépendre de la préservation du secret de l'algorithme. La sécurité ne repose que sur le secret de la clé."

Jusqu'au jour où quelques individus ont remis en cause le principe de Kerckhoffs.

3.2 Le Protocole de Diffie-Hellman (1976)

Alice et Bob vont utiliser une méthode traditionnelle de chiffrement à clé secrète K mais **ils veulent que cette clé ne soit pas acheminée mais déduite par eux seuls par calcul.**

Définition 3 *p premier*
$(\mathbb{Z}_p^*,\ \times)$ est un groupe de $p-1$ éléments
On dit que r est une racine primitive modulo p si :
$\forall x \in \mathbb{Z}_p^*\ \exists \alpha \in \mathbb{N}$ *tel que* $x = r^\alpha$
Autrement dit :
$\mathbb{Z}_p^* \ni n \to r^n \in \mathbb{Z}_p^*$ *appelée exponentiation modulaire est surjective*

Exemples

1. Pour $p = 5$
   2 est une racine primitive modulo 5 car :
   $2^1 = 2, 2^2 = 4, 2^3 = 3, 2^4 = 1$
2. Pour $p = 7$
   2 n'est pas une racine primitive modulo 7 car
   $2^1 = 2,\ 2^2 = 4,\ 2^3 = 1$ et 5 par exemple ne peut pas s'exprimer comme une puissance de 2 dans $\mathbb{Z}_7^*$
   3 est une racine primitive modulo 7 car :
   $3^1 = 3, 3^2 = 2, 3^3 = -1, 3^4 = 4, 3^5 = 5, 3^6 = 1$

Théorème 3 *Pour tout p premier il existe une racine primitive modulo p*

Protocole

1. Alice et Bob se mettent d'accord **publiquement** sur un "grand" nombre premier p (ici on prendra $p = 7$) et une racine primitive modulo p soit r (ici $r = 3$)
2. Alice choisit un nombre a **secret**, connu d'elle seule. Choisissez $a =$ .... (de 1 à 6). Et elle transmet à Bob et à qui veut l'entendre le nombre $\alpha = r^a \bmod p$.
   Que vaut ici $\alpha = 3^a \bmod 7 =$.....?

3. Bob choisit de même un nombre b **secret** et connu de lui seul,et transmet à Alice $\beta = r^b \bmod p$. Choisir une valeur pour b. $b =$ .....
   $\beta =$ .....
4. Ici est le coeur du protocole : **la clé secrète K est** $K = \alpha^b \bmod p = \beta^a \bmod p$
   $\alpha^b \bmod 7 =$ ..... et $\beta^a \bmod 7 =$ .....
   Autrement dit Alice prend ce que lui a envoyé Bob , c'est à dire β et l'élève à la puissance "le nombre secret" d'Alice et obtient la clé secrète K qui servira à chiffrer.
   De même Bob prend ce que lui a envoyé Alice, c'est à dire α et l'élève à la puissance "le nombre secret" de Bob et obtient ô miracle mathématique le même nombre K
5. Oscar a peut-être intercepté les valeurs de p puis r puis r^a puis r^b mais il arrivera difficilement à obtenir a à partir de α ou b à partir de β, parce que l'exponentiation modulaire est une fonction **à sens unique** (voir plus loin)

Théorème 4 *p premier*
r une racine primitive modulo p
Si $a, b \in [\![1, p-1]\!]$ et $\alpha = r^a \ mod\ p$ et $\beta = r^b \ mod\ p$
Alors $\alpha^b = \beta^a \ mod\ p$

Preuve
$\alpha^b = (r^a)^b = r^{ab} = (r^b)^a = \beta^a$

Théorème 5 *(Théorème du logarithme discret)*
r une racine primitive modulo p
$\mathbb{Z}_p^* \ni n \to r^n \in \mathbb{Z}_p^*$ est injective

Preuve
Si $r^n = r^m$ or on a vu sur des exemples plus haut que la suite des puissances de r^k est périodique de période $p-1$ car l'exponentiation modulaire est surjective et à cause du Théorème de Fermat $r^{p-1} \equiv 1\ [p]$
or n et m appartiennent à $[\![1, p-1]\!]$ donc $n = m$
Etant injective et surjective, l'exponentiation modulaire est **bijective** donc inversible

Définition 4 *La fonction réciproque de l'exponentiation modulaire de base r modulo p premier est :*
le logarithme discret de base r, noté $\log_r$*, où r est une racine primitive modulo p premier et définie par :*
$\log_r(r^n) = n$ *avec* $n \in \mathbb{Z}_p^*$

Pratiquement il se trouve que calculer l'image d'un élément de $\mathbb{Z}_p^*$ par l'exponentiation modulaire est "facile" même si p est un "grand" nombre premier (voir exercice)
Par contre pour calculer l'image d'un élément de $\mathbb{Z}_p^*$ par la fonction réciproque, le logarithme discret de base r, le temps mis sera "très grand" (plusieurs années)
On dit alors que l'exponentiation modulaire est une **fonction à sens unique** dans le sens où étant donné une image il est difficile de trouver l'antécédent

3.3 Le cryptosystème RSA (Rivest, Shamir, Adleman) (1977). Chiffrement asymétrique et à clé publique

Chiffrement RSA

1. Alice et Bob ont chacun leurs clés publiques P_A et P_B (comme des numéros de téléphone) et leurs clés secrètes S_A et S_B. Ces clés sont des **paires** d'entiers.
2. Les clés d'Alice (et de même pour Bob) définissent des fonctions $P_A()$ et $S_A()$ à sens unique, **commutatives** c'est à dire :

$$S_A(P_A(M)) = P_A(S_A(M)) = M$$

3. Supposons que Bob veuille envoyer un message M chiffré à Alice. Il se procure la clé publique d'Alice P_A comme un numéro de téléphone dans un annuaire.
4. Bob calcule le texte chiffré $C = P_A(M)$ et envoie C à Alice.
5. Lorsque Alice reçoit C elle calcule $S_A(C) = S_A(P_A(M)) = M$ et elle a accès au texte en clair M.

Comment être sûr qu'un message vient bien de la personne censée l'avoir écrit ?
RSA permet de définir la : **Signature numérique d'un message**

1. Alice envoie un message M' à X . Elle associe à M' une signature définie par $s = S_A(M')$
2. Lorsque X reçoit (M',s) pour s'assurer que c'est bien Alice qui a envoyé M' ,il calcule $P_A(s)$ et regarde s'il obtient M'
3. Si $s = S_A(M')$ alors $P_A(s) = P_A(S_A(M')) = M'$, sinon si $s \neq S_A(M')$ il n'obtiendra pas M'

Un message peut être à la fois **chiffré** et **signé** (voir exercice)

Comment ça marche mathématiquement ?Pour bien comprendre comment fonctionne RSA il nous faut préciser quelques notions mathématiques :

Définition 5 *Attention n un entier non premier on note $\mathbb{Z}_n^*$ l'ensemble des entiers strictement positifs et* ***premiers avec*** *n C'est aussi les éléments inversibles modulo n*

Le nombre d'éléments de $\mathbb{Z}_n^*$ est noté $\phi(n)$ où la fonction ϕ est appelée indicatrice d'Euler

Théorème 6 *1. Si p premier alors $\phi(p) = p - 1$*

2. Si n et m sont premiers entre eux alors $\phi(nm) = \phi(n)\phi(m)$
(Voir le théorème des restes chinois (plus loin))

Théorème 7 *(Théorème d'Euler)*
$\forall a \in \mathbb{Z}_n^*$ *on a* $a^{\phi(n)} \equiv 1\ [n]$

Preuve Soit a un élément quelconque de $\mathbb{Z}_n^*$
Considérons la fonction $f_a : \mathbb{Z}_n^* \ni x \to ax \in \mathbb{Z}_n^*$
injective car si $ax = ax'$ en multipliant à gauche par l'inverse de a on obtient $x = x'$
surjective car pour tout $y \in \mathbb{Z}_n^*$ il existe $x \in \mathbb{Z}_n^*$ tel que $y = ax$

En effet $x = a^{-1}y$

Donc f_a est bijective et on peut voir f_a comme une **permutation** sur l'ensemble **fini** $\mathbb{Z}_n^*$

Donc $\prod_{x\in\mathbb{Z}_n^*} f_a(x) = \prod_{x\in\mathbb{Z}_n^*} x$

Or $\prod_{x\in\mathbb{Z}_n^*} f_a(x) = a^{\phi(n)} \prod_{x\in\mathbb{Z}_n^*} x$ (commutativité de $\times$)

Et donc $a^{\phi(n)} \prod_{x\in\mathbb{Z}_n^*} x = \prod_{x\in\mathbb{Z}_n^*} x$ donc $a^{\phi(n)} \equiv 1\ [n]$

Théorème 8 *(Théorème des restes chinois) Soit n_1 et n_2 deux entiers* ***premiers entre eux*** *et $n = n_1 n_2$ Soit f la fonction définie par :*
$Z_n \ni a \to (a_1, a_2) \in Z_{n_1} \times Z_{n_2}$ *où $a \equiv a_i\ [n_i]$ avec i = 1et 2*

Cette fonction est bijective et respecte l'addition et la multiplication au sens où :
$f(a+b) = f(a) + f(b)$ *et* $f(ab) = f(a) \times f(b)$

Preuve

Montrons l'existence directement de f^{-1}

Puisque n_1 et n_2 sont **premiers entre eux** donc n_1 est inversible dans Z_{n_2} et n_2 est inversible dans Z_{n_1}

Soit $c_1 = n_2 \times (n_2^{-1} \mathrm{mod} n_1)$ on a $c_1 \equiv 1\ [n_1]$ et $c_1 \equiv 0\ [n_2]$
De même $c_2 = n_1 \times (n_1^{-1} \mathrm{mod} n_2)$ on a $c_2 \equiv 1\ [n_2]$ et $c_2 \equiv 0\ [n_1]$

On définit $f^{-1}(a_1, a_2) = (c_1 a_1 + c_2 a_2) \mathrm{mod}\ \ n$

Et $(c_1 a_1 + c_2 a_2) \equiv a_1\ [n_1]$ et $(c_1 a_1 + c_2 a_2) \equiv a_2\ [n_2]$

Exemple

$a \equiv 2\ [5]$ et $a \equiv 3\ [13]$

L'inverse de 13 modulo 5, noté 13_5^{-1} est 2, L'inverse de 5 modulo 13, noté 5_{13}^{-1} est 8

$c_1 = 13 \times 13_5^{-1} = 26$ et $c_2 = 5 \times 5_{13}^{-1} = 40$

Donc $a = 2 \times 26 + 3 \times 40 = 172 \equiv 42\ [65]$

Corollaire 1

Le couple (a, a) de $Z_{n_1} \times Z_{n_2}$ est l'image de a dans Z_n

Corollaire 2

Les inversibles modulo n sont en même quantité que les couples inversibles dans $Z_{n_1} \times Z_{n_2}$ donc $\phi(n) = \phi(n_1) \times \phi(n_2)$

Algorithme RSA

Du côté d'Alice

1. Choisir aléatoirement deux grands nombres premiers p et q différents (de taille chacun de plus de 1024 bits)
2. Calculer $n = pq$
3. Choisir un petit entier e impair premier avec $\phi(n) = (p-1)(q-1)$
4. Calculer d l'inverse de e modulo $\phi(n)$ (Algorithme d'Euclide étendu)
5. $P_A = (e, n)$ est la clé publique RSA d'Alice
6. $S_A = (d, n)$ est la clé secrète RSA d'Alice

Du côté de Bob qui veut envoyer $M \in Z_n$ à Alice :
Le chiffrement est $E(M) = P_A(M) = M^e$ mod n
Déchiffrement du message chiffré $C \in Z_n$ venant de n'importe qui :

Le déchiffrement est $D(C) = S_A(C) = C^d$ mod n

La validité de cet algorithme repose sur le fait que

Pour tout $M \in Z_n$ on doit avoir $P_A(S_A(M) = S_A(P_A(M)) = M$

c'est à dire $M^{ed} \equiv M\ [n]$

Est ce bien le cas ?

Preuve

Puisque e et d sont inverses modulo $\phi(n) = (p-1)(q-1)$ alors il existe k entier tel que

$ed = 1 + k(p-1)(q-1)$

Donc si M n'est pas un multiple de p

$M^{ed} \equiv M(M^{p-1})^{k(q-1)} \equiv M \times (1)^{k(q-1)} \equiv M\ [p]$ (d'après le Théorème de Fermat)

Si $M \equiv 0[p]$ alors $M^{ed} \equiv 0[p]$ et $M^{ed} \equiv M[p]$

De même $M^{ed} \equiv M\ [q]$

Donc $M^{ed} \equiv M\ [p]$ et $M^{ed} \equiv M\ [q]$, d'après le corollaire du théorème des restes chinois

$M^{ed} \equiv M\ [n]$

Sécurité

la sécurité de RSA repose sur la "difficulté" de factoriser des grands entiers, mais cela suppose aussi d'être capable de trouver des grands nombres premiers

Rapidité

On combine RSA avec des systèmes rapides à clé secrète de la manière suivante si Alice veut envoyer un long message M à Bob

1. Elle choisit une clé courte K et chiffre M avec K pour obtenir C
2. Elle chiffre K avec la clé publique RSA de Bob
3. Elle transmet à Bob l'ensemble (C, $P_B(K)$)

4 Qu'est ce qu'un chiffrement "sûr" ?

Comment définir la "sûreté" d'un système cryptographique ?

Nous allons utiliser les idées développées par Claude Shannon dans un article paru en 1948, a Mathematical Theory of Communication"

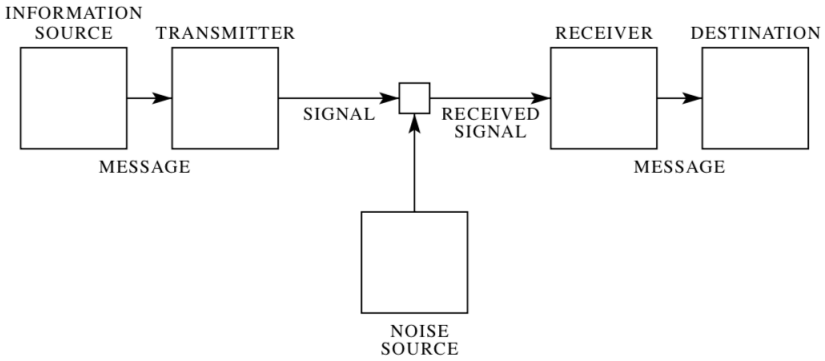


Fig. 1 — Schematic diagram of a general communication system.

Avant cela il nous faut définir une grandeur qui "mesure l'information" dans un message

Définition 6 *Soit X une variable aléatoire **finie** de valeurs x_1, x_2 ...x_n et de probabilités p_1, p_2,..., p_n*

L'entropie de X est $H(X) = \sum_i p_i \ln_2(\frac{1}{p_i})$

Où $\ln_2(x) = \frac{\ln(x)}{\ln(2)}$ *est le logarithme de base 2*

Intuitivement $H(X)$ mesure le degré d'incertitude associé à X

Exemples

1. Un dé équilibré peut être modélisé par une variable aléatoire de valeurs 1,2,...,6 et chaque valeur ayant pour probabilité $\frac{1}{6}$
   Donc $H(X) = \sum \frac{1}{6} \ln_2(6) = \ln_2(6) = 2,585$
2. Un dé truqué peut être modélisé par une variable aléatoire de valeurs 1,2,...,6 où $p_6 = \frac{1}{5}$ et les autres probabilités valant toutes $p_i = \frac{4}{25}$
   Donc $H(X) = \frac{1}{5} \ln_2(5) + 5 \times \frac{4}{25} \ln_2(\frac{25}{4}) = 2,579$

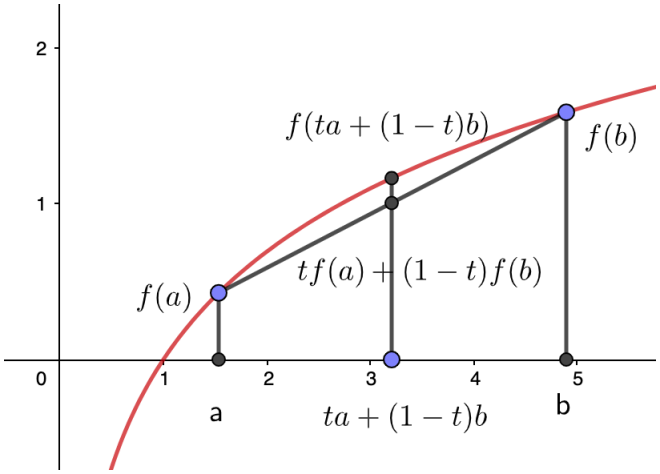
Remarque

Dans la suite on appliquera aussi cette définition à une chaîne de caractères, même si à une chaîne de caractères il existe plusieurs distribution de probabilités possible (voir TP)

Définition 7 *Une fonction f est dite concave sur un intervalle I si*

Pour tout a et b appartenant à I et pour tout $t \in [0,1]$

$tf(a) + (1-t)f(b) \leqslant f(ta + (1-t)b)$



On admettra que la fonction ln est concave sur $]0; +\infty[$

Exercice

1. Justifier géométriquement la concavité de ln
2. Justifier que si f est concave sur I alors pour tout p_i tel que $\sum_i p_i = 1$ pour tout $(a_i)_{1 \leqslant i \leqslant n}$ alors $\sum_i p_i f(a_i) \leqslant \sum_i f(p_i a_i)$ (par récurrence sur $i \geqslant 2$)

Théorème 9 *Pour toute variable aléatoire X ayant n valeurs alors*

$0 \leqslant H(X) \leqslant \ln_2(n)$

Le maximum est atteint pour une variable aléatoire ayant n valeurs équiprobables

Preuve Appliquer le 2) de l'exercice ci-dessus

Définition 8 *Soit X et Y deux variables aléatoires $\Omega fini \longrightarrow \mathbb{R}$*
Une probabilité P est définie sur Ω
X a pour valeurs x_1, x_2 ...x_n et de probabilités p_1, p_2,..., p_n où $p_i = P(X = x_i)$
Y a pour valeurs y_1, y_2 ...y_m et de probabilités q_1, q_2,..., q_m où $q_j = P(Y = y_j)$
***La loi de probabilité conjointe (X, Y)** est définie par*
$P(X = x_i, Y = y_j) = P((X = x_i) \cap (Y = y_j))$
Pour simplifier on remplace $P(X = x_i, Y = y_j)$ par $p(i, j)$
***l'entropie conjointe** $H(X, Y)$ est définie par :*
$H(X,Y) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i,j)})$

Remarque
Comme on l'a vu en TP (Chaîne de Markov) :
$p(i) = \sum_j p(i,j)$ et $p(j) = \sum_i p(i,j)$
$p(i,j) = p(i) \times p_i(j)$ ou $p(i,j) = p(j) \times p_j(i)$

Théorème 10 *Si X et Y sont **indépendantes** (i.e $p(i,j) = p(i) \times p(j)$)*
$H(X,Y) = H(X) + H(Y)$

Théorème 11 $H(X,Y) \leqslant H(X) + H(Y)$

Preuve
(Exercice)

Définition 9 *Soit X et Y deux variables aléatoires de probabilités respectives $(p(i))_{1 \leqslant i \leqslant n}$ et $(p(j))_{1 \leqslant j \leqslant m}$ et de loi de probabilité conjointe $p(i,j)$*
l'entropie conditionnelle de Y relativement à X est définie par :
$H_X(Y) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p_i(j)})$

Théorème 12 $H_X(X) = 0$

Preuve
(Exercice)

Théorème 13 $H(X,Y) = H(X) + H_X(Y) = H(Y) + H_Y(X)$

Preuve
(Exercice)

Théorème 14 $H(Y) \geqslant H_X(Y)$ *avec égalité si X et Y sont indépendantes*

Preuve
Conséquence du théorème 11 et 13

Définition 10 *Un système cryptographique est **inconditionnellement sûr** ou parfait si le message chiffré ne fournit aucune **information** sur le message initial*
En terme d'entropie $H_C(M) = H(M)$ pour tout M et pour tout C

Théorème 15 *Le "one time pad" est un système cryptographique* ***inconditionnellement sûr*** *ou parfait*

Preuve

1. M un message en clair K une clé générée de manière **aléatoire**
   donc on peut considérer M et K comme deux variables aléatoires indépendantes
   donc $H_K(M) = H(M)$
2. $H((M,K),C) = H(M,K) + H_C(M,K)$ or $C = M \oplus K$ donc $H_C(M,K) = 0$
   donc $H((M,K),C) = H(M,K)$
   De même $H((M,C),K = H(M,C)$ et $H((C,K),M) = H(C,K)$ à cause du xor donc
   $H(M,K) = H(M,C) = H(C,K)$
3. Or d'après le théorème 12
   $H(M,C) = H(C) + H_C(M)$ et $H(C,K) = H(C) + H_C(K)$
   or ci-dessus on a montré que $H(M,C) = H(C,K)$
   donc $H_C(M) = H_C(K)$
   de même on montre que $H_K(M) = H_K(C)$
4. D'après le **Théorème 9** si la clef K est vraiment générée de manière aléatoire alors le maximum d'entropie est atteint sur l'ensemble des messages de même longueur n donc
   $H(K) \geqslant H(C)$
   or $H(K,C) = H(K) + H_K(C) = H(C) + H_C(K)$
   donc $0 \leqslant H(K) - H(C) = H_C(K) - H_K(C)$
   donc (3) $H_C(K) = H_C(M) \geqslant H_K(C) = H_K(M)$
   Or (1) $H_K(M) = H(M)$ donc $H_C(M) \geqslant H(M)$
   Mais on a toujours $H(M) \geqslant H_C(M)$
   Donc $H_C(M) = H(M)$

Règles de calcul avec $\sum$

1. **Linéarité de $\sum$**

   (a) $\sum_i (a_i + b_i) = \sum_i a_i + \sum_i b_i$ ressemble à $\int (f(x)+g(x))\mathrm{d}x = \int f(x)\mathrm{d}x + \int g(x)\mathrm{d}x$

   (b) $\sum_i k a_i = k \sum_i a_i$ ressemble à $\int (kf(x))\mathrm{d}x = k \int f(x)\mathrm{d}x$

   On dit "qu'on sort la constante sous le signe somme"

2. **Imaginons que l'on veuille faire la somme de tous les termes d'une matrice a_{ij} de m lignes et n colonnes**

   On peut faire la somme pour chaque ligne i, $s_i = \sum_j a_{ij}$ puis faire la somme des résultats

   $\sum_i s_i = \sum_i (\sum_j a_{ij})$

   Ou alors on fait la somme pour chaque colonne j, $s_j = \sum_i a_{ij}$ puis faire la somme des résultats

   $\sum_j s_j = \sum_j (\sum_i a_{ij})$

   On obtient évidemment le même résultat que l'on note $\sum_{i,j} a_{ij} = \sum_i (\sum_j a_{ij}) = \sum_j (\sum_i a_{ij})$

3. Parfois l'expression a_{ij} peut se décomposer en le produit de deux expressions en i et en j, b_i et c_j

   Par exemple si $a_{ij} = \mathrm{e}^{i+j} = \underbrace{\mathrm{e}^i}_{b_i} \times \underbrace{\mathrm{e}^j}_{c_j}$

   Dans ce cas $\sum_{i,j} a_{ij} = \sum_{i,j} b_i c_j = \sum_i (\sum_j (b_i c_j)) = \sum_i b_i (\sum_j (c_j)) = \sum_j c_j \sum_i b_i$

   Explications

   Dans la troisième expression $\sum_i (\sum_j (b_i c_j))$ si on somme sur j alors b_i est une constante donc on peut la sortir du signe somme, mais alors c'est la somme $\sum_j c_j$ qui est une constante lorqu'on somme sur j donc on peut sortir cette constante du signe somme d'où le produit final

Correction des exercices

1. (Récurence pour la concavité : (hérédité)
   supposons que $\sum_{i=1}^{n} p_i f(a_i) \leqslant f(\sum_{i=1}^{n} p_i a_i)$
   Maintenant $\sum_{i=1}^{n+1} p_i f(a_i)$
   $= p1f(a_1) + ... + p_{n-1}f(a_{n-1}) + (p_n + p_{n+1})(\frac{p_n}{p_n + p_{n+1}} f(a_n) + \frac{p_{n+1}}{p_n + p_{n+1}} f(a_{n+1}))$
   on applique l'hypothèse de récurrence sur les deux derniers termes
   $\leqslant p1f(a_1) + ... + p_{n-1}f(a_{n-1}) + (p_n + p_{n+1}) f(\frac{p_n}{p_n + p_{n+1}} a_n + \frac{p_{n+1}}{p_n + p_{n+1}} a_{n+1})$
   on applique l'hypothèse de récurrence sur les n termes
   $\leqslant f(p_1a_1 + ... + p_{n-1}a_{n-1} + (p_n + p_{n+1})(\frac{p_n}{p_n + p_{n+1}} a_n + \frac{p_{n+1}}{p_n + p_{n+1}} a_{n+1}))$
   $= f(p_1a_1 + ... + p_{n-1}a_{n-1} + p_n a_n + p_{n+1}a_{n+1})$
2. **(Théorème 9)** On applique la concavité de ln pour majorer $H(X)$
   $H(X) = \sum_{i=1}^{n} p_i \ln_2(\frac{1}{p_i}) \leqslant \ln_2(\sum_{i=1}^{n} p_i \times \frac{1}{p_i}) = \ln_2(n)$
3. **(Théorème 10)** Par hypothèse X et Y sont indépendantes donc $p(i,j) = p(i)p(j)$
   Donc $H(X,Y) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i,j)}) = \sum_{i,j} p(i)p(j) \ln_2(\frac{1}{p(i)p(j)})$
   $= \sum_{i,j} [p(i)p(j) \ln_2(\frac{1}{p(i)}) + p(i)p(j) \ln_2(\frac{1}{p(j)})]$
   $= \sum_{i,j} p(i)p(j) \ln_2(\frac{1}{p(i)}) + \sum_{i,j} p(i)p(j) \ln_2(\frac{1}{p(j)})$
   $= \sum_{i} p(i) \ln_2(\frac{1}{p(i)}) \times \sum_{j} p(j) + \sum_{j} p(j) \ln_2(\frac{1}{p(j)}) \times \sum_{i} p(i)$
   $= H(X) + H(Y)$ car $\sum_{j} p(j) = \sum_{i} p(i) = 1$
4. **(Théorème 11)**
   Calculons $H(X) + H(Y) - H(X,Y)$
   $H(X) = \sum_{i} (\sum_{j} p(i,j) \ln_2(\frac{1}{p(i)})) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i)})$
   De même $H(Y) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(j)})$
   Donc $H(X) + H(Y) = \sum_{i,j} p(i,j)(\ln_2(\frac{1}{p(i)}) + \ln_2(\frac{1}{p(j)})) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i)p(j)})$
   Donc $H(X)+H(Y)-H(X,Y) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i)p(j)}) - \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i,j)}) =$
   $\sum_{i,j} p(i,j) \ln_2(\frac{p(i,j)}{p(i)p(j)})$
   Remarque Si f est concave alors $\sum_{i} p_i f(a_i) \leqslant f(\sum_{i} p_i a_i)$
   En multipliant par -1 de part et d'autre

$\sum_i p_i \times (-f(a_i)) \geqslant -f(\sum_i p_i a_i)$ On dit que $-f$ est convexe

Donc $\sum_{i,j} p(i,j) \ln_2(\frac{p(i,j)}{p(i)p(j)}) = \sum_{i,j} p(i,j) \times -\ln_2(\frac{p(i)p(j)}{p(i,j)}) \geqslant \sum_{i,j} -\ln_2(p(i)p(j)) \geqslant 0$

Donc $H(X,Y) \leqslant H(X) + H(Y)$

5. **(Théorème 12)**

   $H_X(Y) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p_i(j)})$ or $p_i(i) = P_{X=x_i}(X = x_i) = 1$

   Donc $H_X(X) = 0$

6. **(Théorème 13)**

   $H_X(Y) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p_i(j)})$ or $p_i(j) = \frac{p(i,j)}{p(i)}$

   Donc $H_X(Y) = \sum_{i,j} p(i,j) \ln_2(\frac{p(i)}{p(i,j)}) = \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i,j)}) - \sum_{i,j} p(i,j) \ln_2(\frac{1}{p(i)}) = H(X,Y) - H(X)$

   Donc $H(X,Y) = H(X) + H_X(Y)$

7. **(Théorème 14)** C'est une conséquence des Théorèmes 11 et 13

   $H(X,Y) = H(X) + H_X(Y) \leqslant H(X) + H(Y)$ donc $H_X(Y) \leqslant H(Y)$