

[sommaire du site](#)

Codage affine

Présentation

Au préalable un petit mot sur la notion de congruence : pour u, v dans Z et n entier ≥ 2 la notation $u \equiv v \pmod{n}$, qui se lit u congru à v modulo n , signifie que $u-v$ est divisible par n ce qui équivaut à dire que u et v ont le même reste lorsqu'on les divise par n ; par exemple $17 \equiv 5 \pmod{3}$ mais aussi $-1 \equiv 1 \pmod{2}$.

On a les résultats suivants très simples :

u, v, r, s étant dans Z et n étant un entier ≥ 2 , si $u \equiv v \pmod{n}$ et si $r \equiv s \pmod{n}$ alors

$u+r \equiv v+s \pmod{n}$; $u-r \equiv v-s \pmod{n}$; $ur \equiv vs \pmod{n}$; $u^p \equiv v^p \pmod{n}$ avec $p \in \mathbb{N}$

et pour tout $k \in Z$ on a $u \equiv v+kn \pmod{n}$

si u et v sont 2 entiers appartenant à $\{0; 1; 2; \dots; n-1\}$ alors $u \equiv v \pmod{n}$ entraîne $u=v$

Notons $E = \{0; 1; 2; \dots; 25\}$

a et b étant 2 entiers choisis dans E , un codage affine consiste après avoir numéroté de 0 à 25 les lettres de l'alphabet à coder une lettre (dite source) de numéro x par la lettre de numéro y , y étant le reste de la division de $ax+b$ par 26. La fonction de codage affine associée aux coefficients a et b est donc la fonction f de E dans E qui à x fait correspondre $f(x)=y$, c'est-à-dire $f(x)$ est le seul élément de l'ensemble $E = \{0; 1; 2; \dots; 25\}$ qui congru à $ax+b$ modulo 26, soit $f(x) \equiv ax+b \pmod{26}$.

Par exemple si $a=17$ et $b=5$ les lettres a, b, c sont codées respectivement par f, w, n . En effet le numéro de a est 0 donc la lettre a est codée par la lettre de numéro $f(0) \equiv 17 \times 0 + 5 = 5$ donc $f(0) = 5$ soit la lettre f ; le numéro de b est 1 donc la lettre b est codée par la lettre de numéro $f(1) \equiv 17 \times 1 + 5 = 22$ donc $f(1) = 22$ soit la lettre w ; le numéro de c est 2 donc la lettre c est codée par la lettre de numéro $f(2) \equiv 17 \times 2 + 5 = 39 \pmod{26}$, et comme $f(2)$ doit être dans E (puisque en fait c'est le reste de la division de 39 par 26) $f(2) = 13$ soit la lettre n ;

En théorie il n'est pas nécessaire de prendre a et b dans E , mais s'ils n'y sont pas, en considérant leurs restes dans la division par 26 on s'y ramène tout de suite : $35x-15 \equiv 9x+11 \pmod{26}$

Jules César (si, si, voir revue Repères n°37 octobre 1999) utilisait un codage affine : $f(x) \equiv x+3 \pmod{26}$; en fait les codages de la forme $f(x) \equiv x+3 \pmod{26}$ reviennent à faire une permutation circulaire sur les lettres. Par exemple pour $b=13$: $a \rightarrow d, b \rightarrow e, c \rightarrow f, \dots, x \rightarrow a, y \rightarrow b, z \rightarrow c$

Une condition évidemment indispensable pour une fonction de codage est que 2 lettres distinctes soient codées de façons différentes, sinon il sera impossible de décoder exactement le message. Dans le cas qui nous intéresse ici cette condition est réalisée si et seulement si a et 26 sont premiers entre eux, c'est-à-dire si a et 26 ont 1 comme seul diviseur positif commun (voir justification dans les parties 2 et 4 de Compléments)

Les seules valeurs de a vérifiant cette condition sont les 12 valeurs suivantes :

1; 3; 5; 7; 9; 11; 15; 17; 19; 21; 23; 25. Dans toute la suite, dès que l'on parlera d'une fonction affine de codage il sera sous-entendu que son a est un de ces 12 nombres

Pour faire immédiatement un codage affine de votre choix :

rentrer le coefficient a (à choisir parmi 1;3;5;7;9;11;15;17;19;21;23;25) dans ce champ

rentrer le coefficient b , entier de 0 (compris) à 25 (compris) dans ce champ

rentrer le message à coder dans le champ ci-dessous : ne mettre que des minuscules non accentuées, les blancs sont permis et votre message peut être plus long que le champ proposé mais ne pas dépasser 500 caractères : si ça dépasse une alerte le signalera et il faudra juste effacer une partie de votre message. En mettant le curseur dans le champ contenant le message codé, via édition, tout sélectionner, puis copier puis coller vous pouvez mettre o vous voulez tout le message codé.

mettre ici le message à coder

Cliquer sur ce bouton pour lancer le codage, si bien sûr Java Script est activé

le résultat va s'afficher ici

Compléments

1) Nombre de codages affines possibles

Tout d'abord $f_{a,b}(x) \equiv ax+b \pmod{26}$ et $f_{a',b'}(x) \equiv a'x+b' \pmod{26}$ seront des fonctions de codage égales (c'est-à-dire $f_{a,b}(x) \equiv f_{a',b'}(x)$ pour tout x dans E) si et seulement si $a \equiv a' \pmod{26}$ et $b \equiv b' \pmod{26}$. Pour le voir il suffit de dire qu'elles sont égales pour $x=0$ et pour $x=1$.

Donc il y a $12 \times 26 = 312$ fonctions affines de codage.

2) Décodage d'une fonction de codage affine

C'est bien beau d'envoyer un message codé, mais le réceptionnaire doit pouvoir décoder facilement le message reçu.

Il faut donc déterminer la fonction de décodage.

Puisque a est premier avec 26, d'après le fameux théorème de Bezout (ou Bézout, selon les auteurs..) il existe u et v dans \mathbb{Z} tels que $au+26v=1$ et donc $au \equiv 1 \pmod{26}$; et ainsi a' est l'élément de E tel que $a' \equiv u \pmod{26}$ (cad $a' = \text{reste de la division de } u \text{ par } 26$)

oici pour tout élément a de E l'élément a' de E tel que $a'a \equiv 1 \pmod{26}$; en fait tous les éléments de E correspondent aux seuls éléments.... inversibles de $\mathbb{Z}/26\mathbb{Z}$.

a	1	3	5	7	9	11	15	17	19	21	23	25
a'	1	9	21	15	3	19	7	23	11	5	17	25

Déterminons maintenant la fonction de décodage de $f_{a,b}(x) \equiv ax+b \pmod{26}$. Cela revient à déterminer pour tout y dans E l'élément x dans E tel que $f_{a,b}(x) \equiv y \pmod{26}$ soit $ax+b \equiv y \pmod{26}$. En multipliant des 2 côtés par a' on obtient l'équation équivalente $x \equiv a'y - a'b \pmod{26}$, ce qui détermine x dans E de façon unique (c'est le reste de la division de $a'y - a'b$ par 26). La fonction de codage $f_{a,b}$ est ainsi une bijection de E dans E et sa fonction de décodage est sa fonction réciproque : c'est la fonction qui à y dans E fait correspondre l'unique élément de E qui est congru à $a'y - a'b$ modulo 26 ; bien entendu on peut remplacer $-a'b$ par l'élément de E tel que $b' \equiv -a'b \pmod{26}$. Concluons :

Pour $a \in E$ et premier avec 26 la fonction affine de codage $f_{a,b}$ est une bijection de E dans E et elle est décodée par la fonction affine $(f_{a,b})^{-1} = f_{a',b'}$ avec a' et b' les seuls éléments de E tels que $aa' \equiv 1 \pmod{26}$ et $b' \equiv -a'b \pmod{26}$

Par exemple pour décoder la fonction $f_{1,13}(x) \equiv x+13 \pmod{26}$ on a $a=1$, $b=13$ donc $a'=1$ et $b' \equiv -a'b \equiv -13 \equiv 13 \pmod{26}$: la fonction de décodage est donc la fonction de codage elle-même, ... ce qui pouvait se prévoir car ajouter 13 puis ajouter 13 c'est ajouter 26 soit 0 modulo 26 : on revient au point de départ.

3) Recherche de toutes les fonctions de codage involutives

On cherche les fonctions de codage telles que $f_{a,b}(f_{a,b}(x)) \equiv x$ pour tout x dans E (c'est-à-dire $(f_{a,b})^{-1} = f_{a,b}$) : on doit donc avoir pour tout x dans E $a(ax+b)+b \equiv x \pmod{26}$ soit $a^2x+ab+b \equiv x \pmod{26}$. Donc (on fait $x=0$ et $x=1$) il faut que $a^2 \equiv 1 \pmod{26}$ et $ab+b \equiv 0 \pmod{26}$: comme on cherche a dans E et 1er avec 26 la table du 2) ci-dessus montre qu'il n'y a que deux possibilités pour a qui sont $a=1$ ou $a=25$.

Si $a=1$ alors $2b \equiv 0 \pmod{26}$ et donc $b=13$ (puisque b est dans E) et on obtient $f_{1,13}(x) \equiv x+13 \pmod{26}$ qui est bien involutive. On peut trouver ce codage dans les lecteurs de news sous le nom ROT13.

Si $a=25$ alors $25b+b=26b$ qui est toujours congru à 0 modulo 26 et on obtient 25 fonctions $f_{25,b}(x) \equiv 25x+b \equiv -x+b \pmod{26}$, pour b quelconque dans E ; il est facile de vérifier que ce sont effectivement des involutions.

4) ue se passe t-il si a n'est pas premier avec 26

Dans ce cas a et 26 ont un diviseur commun $d \geq 2$. On peut alors écrire $26=du$ et $a=dv$; comme $u \perp 26$ on peut trouver x et x' dans E tels que $x'-x=u$, donc $(x'-x)a=udv=26v \equiv 0 \pmod{26}$ et ainsi $ax+b \equiv ax'+b \pmod{26}$, c'est-à-dire x et x' sont codés de la même façon : $f_{a,b}$ n'est pas une bijection de E dans E .

En fait ici ce diviseur commun d à 26 et a ne peut être que 2 ou 13 (puisque $a \leq 25$) donc soit a est pair, soit $a=13$:

si $a=2p$ avec $p \in \{1;2;3;4;5;6;7;7;9;10;11;12\}$ alors on aura $f(x)=f(x')$ si et seulement si $2p(x-x') \equiv 0 \pmod{26}$ soit $p(x-x')$ divisible par 13, et comme p est 1er et $p \leq 12$, $x-x'$ doit être divisible par 13 : je laisse au lecteur le soin de vérifier que dans ce cas $f(E)$ est constitué de 13 éléments : les restes de la division par 13 des nombres : $b, b+1, \dots, b+12$.

si $a=13$ alors $f(x)=f(x')$ si et seulement si $x-x'$ est pair et $f(E)$ est constitué de 2 éléments : b et le reste de la division par 26 de $13+b$, qui sera évidemment $13+b$ si $13+b \leq 25$.

5) "Faiblesse" du chiffrement affine

Si on reçoit un message codé par une fonction affine (bijective bien sr) comme ci-dessus, mais inconnue, est-ce qu'on pourra arriver à déchiffrer "facilement" le message

Aucun problème : la fonction de décodage étant une fonction affine il suffit de passer le message à la "moulinette" des 312 fonctions affines de codage possibles (voir le 1) ci-dessus) et l'une le décodera

On peut s'économiser un peu. En effet, en français, la lettre la plus utilisée est e (17%), puis s (2%), puis n (7,6)..., d'après la revue Pour la Science janvier 2002. Si le message est suffisamment long il y a donc des chances que les lettres qui le composent respectent ces statistiques : il n'est donc pas impossible de repérer par quelles lettres sont codées le e et le s (d'autant plus que la lettre s a la particularité de terminer beaucoup de ...pluriels).

Considérons le message suivant : "stnl nxatq saptq hgnx snv etdunv xplacnv".

On constate que 3 mots se terminent par v, donc on peut penser que s (1) a été codé par v (21) ; et la lettre la plus fréquente du message est n, donc on peut penser que e (4) a été codée par n (13). Pour décoder tout le message il faut donc trouver a' et b' tels que $21a'+b' \equiv 1 \pmod{26}$ (26) et $13a'+b' \equiv 4 \pmod{26}$. Par différence on obtient $a' \equiv 14 \pmod{26}$, mais a' doit être 1er avec 26 et donc on peut se contenter d'essayer les 12 valeurs possibles et on trouve comme seule possibilité $a'=5$; en reportant dans la 1ère équation on obtient $b' \equiv -7 \pmod{26}$ et comme on veut b' dans E (mais à vrai dire ce n'est pas obligé, voir début de la présentation) on prend $b' = -7 + 4 \cdot 26 = 17$, valeur qui vérifie bien la 2ième équation : la fonction de décodage est donc $f_{5,17}(x) \equiv 5x + 17 \pmod{26}$

Pour décoder effectivement tout le message il suffit d'utiliser le programme en début de cette page ; je laisse le soin au lecteur de vérifier que j'ai codé la phrase qu'il vient de découvrir par $f_{21,7}$ ($f_{5,17}$ et $f_{21,7}$ sont réciproques l'une de l'autre).

Bien entendu il existe une méthode permettant de déterminer a' sans essayer les 12 possibilités. En effet l'équation $a' \equiv 14 \pmod{26}$ s'écrit $a' - 26k = 14$ avec k dans Z (et qui se simplifie en $4a' - 13k = 7$, soit $4a' \equiv 7 \pmod{13}$) et on sait résoudre toute équation diophantienne de la forme $ux + vy = w$ avec u, v, w, x, y dans Z et x et y étant les inconnues : voir le paragraphe suivant.

Ce qui précède montre donc que si on sait qu'un message est codé par UNE fonction affine, il est relativement aisé de la trouver et de décoder le message : cependant on peut faire uniquement avec DES fonctions affines des codages beaucoup plus délicats à déchiffrer

oici l'idée de Leon Battista Alberti (1404-1472 ; source Science et ie Junior HS n°53, Juillet 2003) : après avoir codé 5 lettres (par exemple) du message, on code les 5 lettres suivantes avec une autre fonction affine, puis les 5 lettres suivantes avec une troisième fonction affine, et ensuite on réutilise la première fonction affine pour les 5 lettres suivantes, etc...

Si on sait uniquement que le message a été codé par une ou des fonctions affines, mais si on ne sait pas tous les combien on change de fonction affine (ici toutes les 5 lettres) et si on ne connaît pas le nombre de fonctions affines utilisées (ici 3), alors cela devient beaucoup plus délicat à décrypter (les fréquences des lettres ne permettront plus une analyse aussi facile qu'avec une seule fonction affine).

6) Résolution de l'équation diophantienne $ux + vy = w$

Soit $d = \text{pgcd}(u, v)$

Si d ne divise pas w c'est fini pas de solution

Si d divise w l'équation équivaut à $u'x + v'y = w'$ avec u', v', w' les quotients de u, v, w par d .

On recherche une solution particulière ; pour cela on remarque que u' et v' sont 1er entre eux, donc d'après Bezout, il existe p et q dans \mathbb{Z} tels que $u'p + v'q = 1$ (si u et v ne sont pas "évidents" à "voir" on les détermine à l'aide de l'algorithme d'Euclide) donc $u'pw' + v'qw' = w'$ et en faisant la différence membre à membre on obtient $u'(x - pw') + v'(y - qw') = 0$ soit $u'(x - pw') = v'(qw' - y)$. Mais u' est 1er avec v' , donc d'après le théorème de Gauss u' divise $qw' - y$ et il existe k dans \mathbb{Z} tel que $qw' - y = ku'$ soit $y = qw' - ku'$ et donc $x = pw' + kv'$. Réciproquement il est facile de vérifier que les couples $(x, y) = (pw' + kv', qw' - ku')$, pour k quelconque dans \mathbb{Z} , sont effectivement solutions de l'équation ; ce sont donc toutes les solutions.

7) Une conclusion

Il n'aura échappé à aucun élève de TS, bien conditionné :-), que dans la liste précédente de noms célèbres (Euclide, Bezout, Gauss) il manque Fermat. En fait Fermat apparaît, joliment, dans les codages reposant sur des puissances d'entiers. Par exemple le codage RSA qui a été mis au point dans les années 1970 par Rivest, Shamir, Adleman : il est pratiquement impossible à casser. Mais c'est une autre histoire....
