Extension
LibCryptooo

Qu'est-ce que c'est ?

LibCryptooo est une extension de cryptographie élémentaire pour Apache OpenOffice (AOO) ou LibreOffice (LO).
Développée dans un but uniquement pédagogique, *LibCryptooo* permet d'appliquer à tout document texte quelques méthodes de chiffrement classiques et historiques et d'étudier leur déchiffrement. Sans intérêt pratique pour votre correspondance privée (les méthodes utilisées sont dépassées depuis plusieurs siècles), les algorithmes de chiffrement proposés sont un bon prétexte à l'étude des statistiques, mais aussi des fonctions affines ou de l'arithmétique : modulo, diviseurs et nombres premiers entre eux.

Il est indispensable de remarquer que l'utilisation directe de l'extension, sans l'appui d'une activité pédagogique construite, n'apportera rien d'autre qu'un petit tour divertissant dans le monde de la cryptographie.
Cette extension est normalement distribuée avec le dossier *Quelques activités* proposant quelques utilisations effectives de *LibCryptooo* en classe. Un second dossier, *Icônes LibCryptooo*, contient les icônes utilisées par la barre d'outils afin que chacun puisse les insérer dans la consigne de ses propres activités.

Comment installer l'extension ?

Si l'icône de l'extension est sous vos yeux, vous pouvez simplement la double cliquer, l'installation est alors automatique (après validation de votre part).
Si vous avez ouvert LO (version 4.0 ou supérieure) ou AOO (version 4.0 ou supérieure), utilisez le menu *Outils>Gestionnaire des extensions* et cliquez sur le bouton *Ajouter* pour rechercher l'extension sur votre disque dur.
Remarque : il peut être nécessaire d'abaisser le niveau de sécurité de votre logiciel pour la durée de l'installation. Pour ce faire, utilisez le menu *Outils>Options>Sécurité* et réglez le paramétrage de *Sécurité des macros* d'AOO ou de LO sur Faible. Dès l'installation terminée, relevez le niveau de sécurité.
Après l'installation, fermez votre logiciel puis lancez-le à nouveau et ouvrez un document texte.
Une nouvelle barre d'outils (série d'icônes dont la première est un **A**) est disponible, vous pouvez la déplacer, par exemple pour la placer verticalement à droite de votre fenêtre. Si cette barre n'apparaît pas, utilisez le menu *Affichage>Barres d'outils* et cochez la barre *Cryptographie*.

Description et utilisation des icônes

Chacune des icônes de la barre exécute une macro (un petit programme) traitant d'une façon ou d'une autre le texte de la fenêtre active. La fenêtre active doit donc contenir un texte qui vient d'être saisi ou issu de l'ouverture d'un fichier texte pour que les icônes aient une action.
Limitation : le document doit comporter moins de 65 000 caractères.
Notez que l'exécution d'une macro peut être longue pour des textes de plus d'une page.
Le résultat du traitement est en général placé dans une nouvelle fenêtre.

Nettoyage du texte

Conçue dans un but pédagogique, *LibCryptooo* effectue le chiffrement, l'analyse et le déchiffrement de textes simplifiés constitués de caractères majuscules uniquement, de A à Z. L'icône de nettoyage permet d'obtenir un tel texte à partir d'un document typographié de façon quelconque.
Bien que les espaces entre les mots n'aient aucun intérêt pour le chiffrement d'un message secret (et représentent un grave danger en cas d'interception car ils facilitent grandement le décryptage du message), on pourra choisir de supprimer ou non ces espaces. Laisser les espaces permet aux élèves de mieux voir ce qui se passe.

Exemple :

Texte saisi	Texte après nettoyage et suppression des espaces
Il était une fois, une douce et noble princesse armée d'un...	ILETAITUNEFOISUNEDOUCEETNOBLEPRINCESSEARMEEDUN

Le nettoyage du texte est une première étape **obligatoire** avant l'application de l'un des algorithmes de chiffrement proposés ci-dessous.

Méthodes de chiffrement

Chiffre par transposition

Chiffrer un texte par transposition, c'est réorganiser l'écriture du texte en déplaçant les lettres qui le composent d'une façon prédéterminée (une clé) que l'émetteur et le récepteur du message doivent connaître.
Pour un message de longueur *n*, il y a *n!* façons théoriques d'obtenir une transposition du message. Cela devrait conférer à cette méthode un haut degré de sécurité mais en pratique seules des transpositions simples peuvent être réalisées en un temps raisonnable et expliquées au récepteur.
Utilisées depuis le V^{e} siècle avant J-C (la scytale spartiate), les méthodes employées reposent toujours sur un procédé astucieux permettant de transposer rapidement un texte.
La méthode retenue dans *LibCryptooo* est une transposition par grille. Une fois définie la largeur de la grille utilisée (ce nombre constitue la clé), le texte est écrit dans cette grille, caractère par caractère, puis on tourne la grille de 90° dans le sens indirect. On obtient ainsi une transposition du texte initial.

Exemple : avec une largeur clé de 6

Texte en clair	Texte chiffré
ATTAQUEZ DEMAIN A SEIZE HEURES	HSAEAEEIZTUIN TRZ DAEEAEQS MU

Grille de codage correspondante :

A	T	T	A	Q	U
E	Z		D	E	M
A	I	N		A	
S	E	I	Z	E	
H	E	U	R	E	S

Après rotation de 90° :

H	S	A	E	A
E	E	I	Z	T
U	I	N		T
R	Z		D	A
E	E	A	E	Q
S			M	U

Remarque : *LibCryptooo* permet de choisir une grille rectangulaire ou carrée. Choisir une grille carrée permet d'adresser des messages sans préciser la clé, celle-ci se déterminant alors d'après la racine carrée de la longueur du message.

Chiffre par substitution

La méthode par substitution chiffre un message en remplaçant chaque lettre de l'alphabet par une autre préalablement définie par la clé. Cette clé est une chaîne de 26 caractères indiquant les substitutions à effectuer à l'alphabet normalement ordonné.

Exemple : avec la clé ZYXWVUTSRQPONMLKJIHGFEDCBA qui inverse l'ordre alphabétique

Texte en clair	Texte chiffré
ATTAQUEZ DEMAIN A SEIZE HEURES	ZGGZJFVA WVNZRM Z HVRAV SVFIVH

La chaîne saisie doit obligatoirement être constituée des 26 lettres de l'alphabet pour que le déchiffrement soit possible.
Remarque, si l'on ne souhaite réaliser que l'interversion de deux lettres, on peut saisir uniquement ces lettres comme chaîne de substitution.

Exemple : avec la clé AE qui échange les lettres A et E

Texte en clair	Texte chiffré
ATTAQUEZ DEMAIN A SEIZE HEURES	ETTEQUAZ DAMEIN E SAIZA HAURAS

Bien que le nombre de clés possibles soit très important ($26! \approx 4\times10^{26}$), le chiffre par substitution ne constitue pas une méthode cryptographique sure. Si le message est intercepté, il sera très facile de le décrypter par une *analyse de fréquence*. D'autre part la transmission de la clé n'est pas très pratique.
Bien sûr, si le message est très court comme dans l'exemple ci-dessus et qu'à chaque message la clé est modifiée,

cette méthode gagne beaucoup en sûreté. Mais en pratique états et armées ont besoin d'envoyer de très nombreux messages (qui même très courts font un ensemble très long) et générer de très nombreuses clés différentes pose un problème technique évident. Comment les retenir ? Comment savoir quelle clé a été utilisée pour le chiffrement ?

Chiffre de César

Méthode de chiffrement plus simple encore que la substitution générale, le chiffre de César doit son nom à Jules César dont on sait qu'il employa cette méthode pour sa correspondance. Son grand avantage tient justement dans sa facilité d'utilisation, même si le déchiffrement d'un message est d'une simplicité déconcertante.
Pour chiffrer un message, on décale l'alphabet d'un certain nombre de rangs préalablement défini par la clé. Une fois le Z atteint, on recommence un nouvel alphabet à partir du A.
Par convention, une clé positive correspond à un décalage à droite, une clé négative à un décalage à gauche.
Exemple : avec une clé de 3, un A sera chiffré en D, un B deviendra un E … un Z deviendra un C.

Exemple : avec la clé 13

Texte en clair	Texte chiffré
ATTAQUEZ DEMAIN A SEIZE HEURES	NGGNDHRM QRZNVA N FRVMR URHERF

Le code est très peu sûr : il n'y a que 26 clés possibles.

Chiffre affine

Le chiffre affine est une généralisation du code de César qui nécessite bien des calculs et n'a pas d'importance historique mais il possède d'intéressantes propriétés mathématiques.
Soient a et b deux nombres entiers compris entre 0 et 25, $f(x) = ax + b$ définit une fonction affine.
En associant A au nombre 0 puis à chaque lettre de l'alphabet son numéro d'ordre, on peut calculer l'image de chacune (mod 26) puis coder le message.
Prenons $f(x) = 3x + 2$, on a $f(0) = 2 \rightarrow$ A devient C, $f(1) = 5 \rightarrow$ B devient F, $f(2) = 8 \rightarrow$ C devient I, … ,
$f(25) = 77 \equiv 25\ (26) \rightarrow$ Z devient Z.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
C	F	I	L	O	R	U	X	A	D	G	J	M	P	S	V	Y	B	E	H	K	N	Q	T	W	Z

Exemple : avec la clé a = 3 et b = 2

Texte en clair	Texte chiffré
ATTAQUEZ DEMAIN A SEIZE HEURES	CHHCYKOZ LOMCAP C EOAZO XOKBOE

Il faut remarquer que seuls les nombres a premiers avec 26 permettent de produire un message codé déchiffrable (sinon plusieurs lettres seront codées de la même manière sans que l'on puisse les différencier grâce à la clé).
D'autre part l'arithmétique modulo 26 permet d'associer à la fonction de chiffrement f une fonction inverse de déchiffrement f^{-1} qui est aussi affine.
Bien que cela ne fasse mathématiquement aucune différence avec le cas $b>0$, *LibCryptooo* autorise la saisie d'un coefficient b négatif (plus naturel lorsque l'on détermine la fonction inverse).

Exemple : le message précédent, chiffré avec a = 3 et b = 2, sera déchiffré avec a' = 9 et b' = -18

Texte chiffré	Texte déchiffré
CHHCYKOZ LOMCAP C EOAZO XOKBOE	ATTAQUEZ DEMAIN A SEIZE HEURES

Ici encore, même en l'absence de la clé, un message intercepté est attaquable par *analyse de fréquence* et le nombre de clés est très limité : 12 pour a et 26 pour b font 312 clés seulement.

Chiffre de Vigenère

Apparu au tout début de la Renaissance, le chiffre de Vigenère renforce considérablement la sécurité des messages secrets en permettant l'emploi de clés simples et faciles tant à retenir qu'à changer tout en rendant inefficace l'attaque par *analyse de fréquence* directe. Bien qu'assez peu utilisé (parce que moins pratique d'emploi que les codes plus simples), il garantit la sécurité des correspondances secrètes pendant plus de trois siècles, jusqu'à ce que Charles Babbage parvienne à le briser pour le compte de la couronne britannique.
Son principe est basé sur l'utilisation de plusieurs chiffres de César, définis préalablement par un mot clé.
Par exemple, la clé LOVE (soit numériquement 11-14-21-4) indique que la première lettre du message sera décalée de 11 rangs, la deuxième de 14 rangs, la troisième de 21 rangs, la quatrième de 4 rangs (et on recommence pour les lettres suivantes).

Exemple : avec la clé LOVE

Texte en clair	Texte chiffré
ATTAQUEZ DEMAIN A SEIZE HEURES	LHOEBIZD RZQLWI L NITNZ SSPVPG

Grâce à la clé, deux lettres identiques peuvent être chiffrées différemment et deux lettres différentes peuvent être codées de façon identique.

Remarque :
l'algorithme programmé considère les espaces comme occupant la place d'une lettre mais sans être chiffrés.

A	T	T	A	Q	U	E	Z		D	E	M	A	I	N		A		S	E	I	Z	E		H	E	U	R	E	S
L	*O*	*V*	*E*	*L*	*O*	*V*	*E*	*L*	*O*	*V*	*E*	*L*	*O*	*V*	*E*	*L*	*O*	*V*	*E*	*L*	*O*	*V*	*E*	*L*	*O*	*V*	*E*	*L*	*O*
L	H	O	E	B	I	Z	D		R	Z	Q	L	W	I		L		N	I	T	N	Z		S	S	P	V	P	G

Analyse de documents chiffrés

Calculs statistiques

Que le document soit chiffré ou non (mais nettoyé), cliquer sur cette icône ouvre une feuille de calcul et affiche les occurrences de chaque lettre. Charge est laissée à l'élève de calculer le nombre total de lettres ainsi que la fréquence de chaque lettre dans le texte (le format des cellules C2 à C27 pourra être modifié pour obtenir un affichage plus lisible).
Ce travail constitue une initiation à l'*analyse de fréquence*.
Dans les langues européennes, la lettre E se présente avec une fréquence bien supérieure à celle des autres lettres de l'alphabet. Repérer la lettre la plus fréquente d'un texte codé par le chiffre de César permet d'en déduire le décalage effectué et donc de décrypter le message par un décalage opposé.

Analyse de fréquence

Que le document soit chiffré ou non (mais nettoyé), cette icône ouvre une feuille de calcul et effectue le calcul de la fréquence de chaque lettre dans le texte.
Ces fréquences sont ensuite comparées aux fréquences moyennes des lettres dans la langue française.
Enfin, le programme associe dans l'ordre décroissant des fréquences les lettres du texte à celles de l'alphabet et propose une chaîne de substitution.
L'analyse de fréquence permet de décrypter des messages codés par un algorithme de substitution ou une fonction affine dont on ignore la clé.
La correspondance n'est que très rarement parfaite, même pour un texte long (car les fréquences des lettres dans le message ne sont pas exactement égales aux fréquences moyennes, fréquences qui sont parfois très proches les unes des autres et varient selon le niveau de langue et le type de texte). La chaîne de substitution ne décrypte donc que partiellement le message. L'observation minutieuse du résultat permet en général de proposer des interversions de lettres qui améliorent le décryptage.
Pour une meilleure correspondance, il est possible d'analyser les fréquences des doublets de lettres et des triplets. Ce type d'analyse n'est actuellement pas implémenté dans *LibCryptooo*.

Tableau des fréquences (%) moyennes utilisé :

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
8,01	0,88	3,23	3,91	17,5	1,05	1,07	0,88	7,35	0,44	0,05	5,77	2,90	7,22	5,43	2,94	1,14	6,69	8,17	7,07	6,00	1,41	0,02	0,47	0,30	0,12

Soit dans l'ordre décroissant : ESAINTRULODCPMVQGFHBXJYZKW

L'analyse de fréquence permet également de déterminer la langue d'origine d'un texte codé en comparant par exemple son graphique de fréquence à celui de plusieurs langues.

Recherche de clé de Babbage

La méthode de Babbage vise à résoudre le problème posé par le chiffre de Vigenère.
En raison de la périodicité de l'utilisation du mot clé, le chiffre de Vigenère qui ne peut pas être brisé par une simple *analyse de fréquence*, peut être cassé par une analyse multiple pour peu que l'on connaisse la longueur de la clé. La méthode de Babbage débute donc par la détermination de la longueur du mot clé (par la recherche de séquences de lettres répétées à l'identique dans le message) ; ceci n'est pas implémenté dans *LibCryptooo*.
Une fois la longueur *n* de la clé connue, une analyse de fréquence pour l'ensemble des lettres en position 1+*kn* donne le décalage associé et donc la première lettre du mot clé. En procédant de même avec les groupes de lettres en position 2+*kn*, 3+*kn*... on en déduit l'ensemble du mot clé.
L'exactitude du résultat pour un seul message est fortement dépendante de la longueur du message et de celle de la clé.

Déchiffrement

Les dernières icônes de la barre d'outils sont des utilitaires permettant de déchiffrer exactement des messages dont on connaît la clé.

Déchiffrement par transposition avec clé

Détermine la clé de déchiffrement et effectue le codage inverse (rotation de 90° dans le sens direct).
La clé de chiffrement correspondant à la largeur de la grille de codage, la clé de déchiffrement doit être égale à sa hauteur (que le récepteur du message peut calculer d'après la clé de chiffrement et la longueur du message).

Déchiffrement affine avec clé

Détermine la fonction inverse de la clé (ce qui dépasse les programmes de collège mais peut être traité par un algorithme au lycée) et déchiffre le message.

Déchiffrement de Vigenère avec clé

Détermine la clé inverse du mot clé de chiffrement (ce qui n'est pas bien difficile mais sans intérêt) et déchiffre le message.

Notes techniques

● Le fichier d'extension à installer est nommé *LibCryptOoo-A-#* pour sa version Apache OpenOffice et *LibCryptOoo-L-#* pour sa version LibreOffice. Installer obligatoirement la version dédiée à chaque logiciel.

● LibCryptOoo ne peut être installé sur des versions de LO ou AOO inférieures à 4.0.

● Longueur des messages : le texte doit contenir moins de 65 536 caractères, espaces compris.

● Nettoyage du texte : cette macro supprime les accents et passe les caractères en majuscules. Tout autre type de caractère (chiffres, ponctuation, caractères invisibles, symboles divers) est supprimé, sauf les espaces selon l'option choisie par l'utilisateur.
æ et œ sont remplacés par A et O.

● Analyse du nettoyage d'un texte : avant l'exécution d'une macro, les caractères du message sont analysés pour vérifier que le texte ne contient que des caractères compris entre A et Z.
Cette vérification n'est exhaustive que si le texte contient moins de 1000 caractères. Au delà, 1000 caractères tirés au hasard sont testés. Un texte long non nettoyé (par exemple parce qu'il a été saisi directement en majuscules par l'utilisateur) peut donc contenir des caractères inappropriés, sans que cela soit détecté, et provoquer une erreur d'exécution de la macro.

● Temps d'exécution : le temps d'exécution d'une macro est proportionnel à la longueur du message et peut devenir rédhibitoire pour des messages longs (en particulier pour l'algorithme de transposition).

● Transposition : la largeur maximale d'une grille rectangulaire de transposition est fixée à 100 caractères. Ce nombre peut être dépassé pour une grille carrée.

● Vigenère : la longueur du mot clé est limitée à douze caractères.

● Analyse de fréquence : les fréquences moyennes des lettres retenues pour un texte rédigé en français ne prétendent pas à une valeur plus exacte que les nombreuses tables (aux résultats très variés) disponibles dans la littérature sur ce sujet.

● Image « Alice » issue de Wikipedia.

● Contacter l'auteur : vincent.everaert@ac-rouen.fr