TD1 — Stéganographie & Chiffrements historiques

Ex1. Manipulation d'octets et représentations

(a) Soit l'octet `10101001`. Donnez sa représentation héxadécimal.

(b) Quel entier en base décimale correspond à l'octet `10101001` ?

(c) Que vaut `238 & 1` ? Que vaut `239 & 254`? Que vaut `(239 & 254) | 1` ?

Ex2. Stéganographie Nous avons récupéré un extrait d'image (une séquence de pixels RGB). Il est caché dans cette image un mot de trois lettres suivant l'algorithme *LSB* mais en se basant sur l'avant dernier bit. Toutes les composantes couleurs contiennent un seul bit du mot mystère. Pour rappel, un caractère = un octet.

id pixel	0	1	2	3
Pixel	(153, 103, 42)	(226, 213, 52)	(102, 221, 152)	(250, 2, 69)
id pixel	4	5	6	7
Pixel	(66, 242, 162)	(131, 209, 75)	(86, 33, 39)	(236, 36, 126)

Ex3. Chiffre monoalphabétique
Le chiffrement de César consiste à remplacer chaque lettre du message en clair par la lettre située trois places plus loin dans l'alphabet :

```
entree: 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
sortie: d e f g h i j k l m n o p q r s t u v w x y z a b c
```

Dans cet exercice on suppose que la ponctuation et les espaces sont préservées.

(a) Chiffrer le message « `alea jacta est` »

(b) Déchiffrer le message « `yhuflqjhwrula` »

Plus généralement, un chiffrement par décalage utilise une clé secrète k qui donne la portée du décalage (dans le chiffrement de César $k = 3$). Si on code les 26 lettres de l'alphabet par des entiers (a=0, b=1,..., z=25), le message chiffré C est obtenu à partir du message en clair M par la formule $C = E(k, M) = M + k \pmod{26}$.

(c) Déchiffrer le message suivant, en français, sans connaître la clé : « `iuq, mvbmvla-bc tm dwt vwqz lma kwzjmicf acz vwa xtiqvma ?` »

Ex4. Chiffre polyalphabétique Le chiffrement de Vigenère étend le chiffrement par décalage en utilisant une clé composée de plusieurs décalages répétés périodiquement. Dans cet exercice on suppose que la ponctuation et les espaces sont supprimés. La dernière page du sujet comporte un message en français chiffré par le chiffrement de Vigenère (`PFEOH...`) ainsi que différentes statistiques sur des sous-séquences de la forme $C[ax + b]$ extraites de ce message.

(a) Quel est l'indice de coïncidence attendu pour un texte aléatoire suffisamment long ?

(b) Retrouver la longueur probable de la clé en étudiant les indices de coïncidence.

(c) Retrouver la clé et déchiffrer le dernier mot du message.

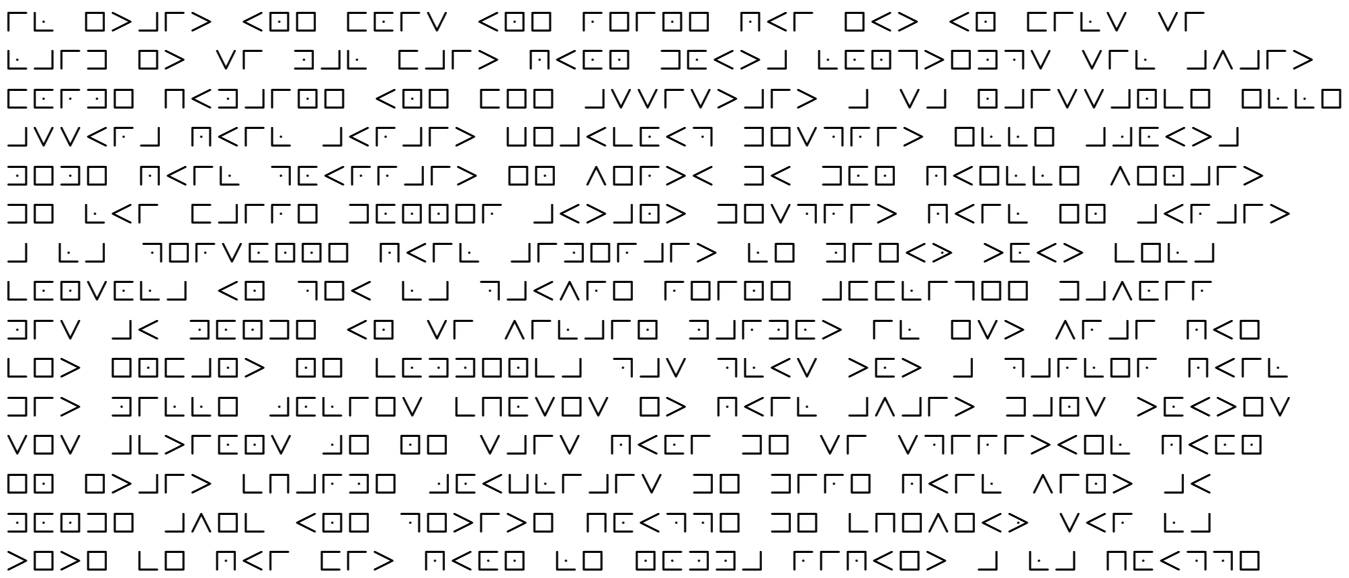
Indices de coïncidence typiques (source : Wikipedia) : russe 0,0529 ; anglais 0,0667 ; français 0,0778 ; allemand 0,0762 ; finnois 0,0737 ; japonais 0,0772.

Ex5. L'oeil qui voit tout L'étrange alignement de graffitis reproduit en fin d'exercice a été découvert sous la peinture sur le mur de la salle E11. Le message semble ancien, il est composé à partir de 22 symboles distincts.

(a) L'utilisation de symboles exotiques rend-elle le chiffrement d'un message plus sûr ?

(b) Calculer l'indice de coïncidence du message. Que peut-on en déduire ?

(c) *(bonus)* Déchiffrer le graffiti suivant découvert sur le même pan de mur :



Nombre d'occurences de chaque symbole : ⌜ : 62 ; ⌞ : 38 ; □ : 82 ; > : 45 ; ⌟ : 61 ; < : 51 ; ⊡ : 39 ; ⊏ : 10 ; ⊑ : 31 ; ⋁ : 37 ; ⌜ : 26 ; ⊓ : 16 ; ⊐ : 17 ; ⊐ : 19 ; ⌝ : 2 ; ⌝ : 17 ; ⋀ : 11 ; ⊓ : 6 ; ⌞ : 13 ; ⊔ : 2 ; ⌟ : 4 ; > : 2.

Ex6. Chiffrement par xor (noté $\oplus$)

(a) M, K, C sont des blocs de bits. Supposons que $\text{long}(M) = 8$ et $\text{long}(K) = 4$. On rappelle que $E(K, M) = M \oplus k$. Chiffrer le message 01011100 avec la clé 0101.

(b) Malheureusement votre adversaire sait que le message clair M commence par 0101, et il sait que $C = 01101100$. Que peut-il faire pour trouver la clé ? Comment s'appelle ce type d'attaque ?

(c) L'attaque précédente a-t-elle encore un intérêt si $\text{long}(K) = \text{long}(M) = 8$, et la clé K n'est plus utilisée par la suite (masque jetable) ?

(d) Maintenant $\text{long}(K) = \text{long}(M) = 4$. On veut comparer $E(K, M) = M \oplus K$ (xor bit à bit), et $E'(K, M) = (M + K) \pmod{16}$ (addition modulo 16 des entiers M et k codés en base 2). Est-ce la même chose ?

	SP	!	"	#	$	%	&	'	(	)	*	+	,	-	.	/	0	1	2
Déc	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
Hex	0x20	0x21	0x22	0x23	0x24	0x25	0x26	0x27	0x28	0x29	0x2A	0x2B	0x2C	0x2D	0x2E	0x2F	0x30	0x31	0x32

	3	4	5	6	7	8	9	:	;	<	=	>	?	@	A	B	C	D	E
Déc	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
Hex	0x33	0x34	0x35	0x36	0x37	0x38	0x39	0x3A	0x3B	0x3C	0x3D	0x3E	0x3F	0x40	0x41	0x42	0x43	0x44	0x45

	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Déc	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
Hex	0x46	0x47	0x48	0x49	0x4A	0x4B	0x4C	0x4D	0x4E	0x4F	0x50	0x51	0x52	0x53	0x54	0x55	0x56	0x57	0x58

	Y	Z	[	\	]	^	_	‘	a	b	c	d	e	f	g	h	i	j	k
Déc	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107
Hex	0x59	0x5A	0x5B	0x5C	0x5D	0x5E	0x5F	0x60	0x61	0x62	0x63	0x64	0x65	0x66	0x67	0x68	0x69	0x6A	0x6B

	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~
Déc	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126
Hex	0x6C	0x6D	0x6E	0x6F	0x70	0x71	0x72	0x73	0x74	0x75	0x76	0x77	0x78	0x79	0x7A	0x7B	0x7C	0x7D	0x7E

Figure 1: Quelques codes ASCII

PFEOH OPCBL JVBCZ FVCYQ LHZEB UXOUN KVBCZ OUMYW EGUWZ YZRPR TSBZP CBLYE BYRSZ JGFPM OSUMY SCYOZ SYXSU NGIIL AWAXK HVOZS ZFKGJ FUQOY YGVHT
OUNGU YUTRL PUZLY JOUMR OALOD SYKBJ YOBAY JSSUI WAYJS SOTWC YXGPN KSAXK ZHPOZ SYISU YYHJY VSUXG BAJGG BHPCB LJCUN RVPMZ CPLKO PNMOY XKGVO
BSUCX EBYRS ZCDXH HBWLL SWSFK EBUZF LWKBA KAOAL KJPHM HKYAL YCKBK YTCAU HZLXG BZFKJ LHKAL HZEBC SSANG WAUOB ZCKBI LGBSY JSZFK AHNOB SYYQS
IIVLM KHSYY PVOXU LIOGK YVOYC YQLHK HHCZB POTOZ MGIAX KDPWG FKMUI KYHCB LMIPA TCUMT WBHKQ OUYGL GKBLY KBWLU QLMYW VHTWB HKFLP UZAYJ SJIRW
LLYRH HYZHP OUUYJ SSUGG UCABL YTHYY KRLHU HYYJW ANXSZ LKRVO ZSZYO UUYAF TITGP YAFSY XCPHO ALGKI UYHSS FKDLH JOPMU BKYRO YLUBZ YZRLF GFYIT
BLMYS ZUROQ OYHPW KRLJG FPM

a	*b*	*IC*	*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*
1	0	0.048	24	34	22	4	7	17	25	33	15	19	35	38	16	9	35	24	7	15	36	13	37	13	17	14	58	31
2	0	0.060	20	17	2	1	2	0	13	10	5	14	34	23	3	0	13	19	1	7	15	13	24	9	1	7	23	23
2	1	0.054	4	17	20	3	5	17	12	23	10	5	1	15	13	9	22	5	6	8	21	0	13	4	16	7	35	8
3	0	0.048	6	12	7	0	1	6	9	13	6	6	10	12	4	2	14	10	1	7	13	5	14	3	6	4	20	9
3	1	0.044	10	7	7	3	3	6	6	9	5	5	15	15	10	4	10	5	2	3	11	5	12	6	8	5	16	11
3	2	0.048	8	15	8	1	3	5	10	11	4	8	10	11	2	3	11	9	4	5	12	3	11	4	3	5	22	11
4	0	0.082	14	15	2	0	1	0	0	7	2	4	6	23	0	0	2	16	1	0	13	1	15	7	1	0	8	12
4	1	0.098	1	0	11	0	0	8	3	15	5	3	1	14	13	9	8	4	0	1	0	0	9	0	3	6	35	1
4	2	0.081	6	2	0	1	1	0	13	3	3	10	28	0	3	0	11	3	0	7	2	12	9	2	0	7	15	11
4	3	0.068	3	17	9	3	5	9	9	8	5	2	0	1	0	0	14	1	6	7	21	0	4	4	13	1	0	7
5	0	0.053	4	7	2	0	3	5	2	9	1	6	10	9	1	1	8	5	1	1	10	3	5	2	3	2	14	6
5	1	0.047	6	6	7	1	0	3	6	8	3	3	8	5	1	2	5	3	1	2	10	2	9	7	3	2	11	6
5	2	0.043	5	6	5	2	2	4	3	6	5	1	6	7	6	2	6	7	1	5	3	5	11	2	2	2	11	5
5	3	0.046	7	8	5	0	1	2	7	6	2	4	4	8	5	1	10	5	1	5	7	0	4	2	4	6	11	4
5	4	0.045	2	7	3	1	1	3	7	4	4	5	7	9	3	3	6	4	3	2	6	3	8	0	5	2	11	10
6	0	0.058	5	6	0	0	0	0	5	3	1	6	10	6	1	0	6	9	0	3	3	5	11	3	1	1	7	8
6	1	0.045	2	5	6	3	2	6	3	6	2	1	1	5	8	4	7	2	2	2	4	0	4	2	8	1	10	4
6	2	0.055	7	9	1	1	1	0	5	4	1	4	10	7	0	0	4	7	1	3	5	3	5	2	0	2	10	8
6	3	0.061	1	6	7	0	1	6	4	10	5	0	0	6	3	2	8	1	1	4	10	0	3	0	5	3	13	1
6	4	0.061	8	2	1	0	1	0	3	3	3	4	14	10	2	0	3	3	0	1	7	5	8	4	0	4	6	7
6	5	0.048	1	6	7	0	2	5	5	7	3	4	0	4	2	3	7	2	3	2	7	0	6	2	3	3	12	3

Chaque ligne de cette table présente le nombre d'occurences de chaque lettre dans la sous-séquence $C[ax + b]$ pour les valeurs a et b présentées dans les deux premières colonnes. La colonne IC indique l'indice de coïncidence de cette sous-séquence.