



MAS 335

Cryptography

Notes 3: Stream ciphers

Spring 2008

FZFGW BOPFW LWKRA SUQSY JHSIJ DHFVW ICCWA YHFRY GMEIJ
XWPXW WCKXZ JPXRC FBASX MOSMF LBLXZ NBDXG ICLRU JCOXO
NQBWZ JVXHH JSMIV NBQSL MSYSG PVBVK NGQIJ BOPVW FRFRY
GIQML MOARG UWZXM WSPSJ HCKZW WGXXA TBPMF NHXRV BVXXA
XHEIM XSLJS GLOL MCRKZ YOIMU JKFXZ TIQTA HHRVW XCOGG
SJBVK FHFSF XGLWZ JKXWU TBPMV JFFRY NBEIJ TKKQA SRXWO
JZIEK XVBGG ZZAJG WHEIZ THAEQ ROAIZ JFCIW QJBVQ XZBIH
DOKHK YIMMV BVBXZ JFQLW UZBEK ZFBSX ROHMF LOAEA XMZLS
NBTSM QRYIO TFQLL MSQVG ZPIIG KUBXL NBDYH FBATA HYFRY
YVBHS NGFIK BVBRK ZRAIF QMXAZ NHBVS GPF XO NHETA SYBCW
XFXRU QCPIT DVBV

a	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
b	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	A
c	C	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
d	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
e	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	D
f	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
g	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
h	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
i	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
j	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
k	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
l	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
m	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
n	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
o	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
p	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
r	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
s	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
t	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
u	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
v	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
w	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
x	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
y	Y	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
z	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

Table 1: Vigenère square

Alice was beginning to get very tired of sitting by her sister on the bank and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, and “what is the use of a book,” thought Alice, “without pictures or conversations?” So she was considering, in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

Step 1 asks us to guess the keyword length. We notice that the digram **HE** occurs at positions 182, 287 and 442 in the ciphertext, and in the first two of these it is part of the trigram **HEI**. So by the first step of Kasiski's algorithm, the keyword length should divide $\text{gcd}(105, 155) = 5$. (I have anticipated this by writing the cipher in blocks of 5; usually Alice will not be so helpful!)

The first of the five substrings that we have to analyse is obtained by taking the first letter of each block; it is **FBLSDIYGXWJFMLNIJNJNMPNBFGMUWHWTNBXXGMYJTHXSFXTJNTSJXZWTRJQXDYBJUZRLXNQTZKMFHYNBZQNGNSXQD**. The letter frequencies in this substring are given in the third column of Table 2.

We calculate the chi-squared values using the frequency data from *Alice's Adventures in Wonderland*. Table 2 gives the calculation for shifts 0 and 5; it is easy to automate this to work out all values. The answers would not be very different if we had used different data for the frequencies.

BV	8	1	[118, 176, 228, 308, 326, 416, 433, 463]
FR	5	1	[38, 131, 133, 253, 403]
NB	5	15	[76, 106, 256, 361, 391]
RY	5	1	[39, 134, 254, 367, 404]
VB	8	1	[117, 175, 277, 325, 327, 407, 417, 462]
XZ	5	1	[54, 74, 209, 311, 329]
ZJ	5	5	[55, 95, 240, 300, 330]

BV	8	1	[er, wh, er, er, wh, wh, er, er]
FR	5	1	[in, ad, in, in, in]
NB	5	15	[in, in, in, in, in]
RY	5	1	[ng, ng, ng, db, ng]
VB	8	1	[he, dw, he, dw, he, he, he, he]
XZ	5	1	[th, th, th, sl, th]
ZJ	5	5	[he, he, he, he, he]

FRY	4	5	[38, 133, 253, 403]
			[ing, ing, ing, ing]
VBV	4	1	[117, 175, 325, 462]
			[her, dwh, dwh, her]

Letter	Frequency %	Observed	Expected Shift 0	Expected Shift 5
A	8.15	0	7.58	0.73
B	1.37	5	1.27	2.32
C	2.21	0	2.06	0.12
D	4.58	3	4.26	1.96
E	12.61	0	11.73	0.65
F	1.86	5	1.73	7.58
G	2.36	4	2.20	1.27
H	6.85	3	6.37	2.06
I	6.97	2	6.48	4.26
J	0.14	11	0.13	11.73
K	1.07	1	1.00	1.73
L	4.37	3	4.06	2.20
M	1.96	5	1.82	6.37
N	6.52	11	6.06	6.48
O	7.58	0	7.05	0.13
P	1.40	1	1.30	1.00
Q	0.19	4	0.18	4.06
R	5.02	2	4.67	1.82
S	6.05	4	5.63	6.06
T	9.93	6	9.23	7.05
U	3.22	2	2.99	1.30
V	0.78	0	0.73	0.18
W	2.49	4	2.32	4.67
X	0.13	9	0.12	5.63
Y	2.11	4	1.96	9.23
Z	0.07	4	0.65	2.99
$\sum(o - e)^2/e$			1949.79	23.99

Table 2: A chi-squared calculation

AS	4	2	[15, 63, 265, 445]
BV	8	1	[118, 176, 228, 308, 326, 416, 433, 463]
EI	4	5	[43, 183, 258, 288]
FB	3	1	[61, 342, 396]
FQ	3	1	[332, 372, 425]
FR	5	1	[38, 131, 133, 253, 403]
FX	4	1	[208, 235, 438, 452]
HE	3	5	[182, 287, 442]
HF	4	1	[27, 37, 232, 395]
IJ	4	5	[24, 44, 124, 259]
IM	3	1	[184, 203, 322]
JB	3	2	[125, 227, 307]
JF	3	10	[251, 301, 331]
KX	3	3	[53, 242, 275]
KZ	4	1	[158, 199, 340, 420]
LM	4	5	[110, 140, 195, 375]
LW	3	1	[11, 238, 334]
MF	3	20	[69, 169, 349]
NB	5	15	[76, 106, 256, 361, 391]
NH	3	10	[171, 431, 441]
OA	3	5	[142, 297, 352]
RY	5	1	[39, 134, 254, 367, 404]
SG	3	1	[114, 190, 435]
SM	3	2	[68, 102, 364]
SY	3	1	[19, 112, 446]
TA	3	5	[214, 399, 444]
VB	8	1	[117, 175, 277, 325, 327, 407, 417, 462]
VW	3	10	[29, 129, 219]
WZ	3	1	[94, 147, 239]
XA	3	3	[164, 179, 428]
XM	3	3	[65, 149, 356]
XR	4	1	[58, 173, 345, 453]
XW	4	1	[46, 49, 243, 268]
XZ	5	1	[54, 74, 209, 311, 329]
ZJ	5	5	[55, 95, 240, 300, 330]

AS	4	2	[in, do, in, in]
BV	8	1	[er, wh, er, er, wh, wh, er, er]
EI	4	5	[he, he, he, he]
FB	3	1	[an, re, an]
FQ	3	1	[rt, rt, nl]
FR	5	1	[in, ad, in, in, in]
FX	4	1	[it, ns, it, ra]
HE	3	5	[th, th, th]
HF	4	1	[ti, ti, ti, pa]
IJ	4	5	[er, er, er, er]
IM	3	1	[eu, li, up]
JB	3	2	[rw, ve, ve]
JF	3	10	[er, er, er]
KX	3	3	[nt, wa, ss]
KZ	4	1	[nv, gh, su, su]
LM	4	5	[th, th, th, th]
LW	3	1	[gi, os, he]
MF	3	20	[in, in, in]
NB	5	15	[in, in, in, in, in]
NH	3	10	[it, it, it]
OA	3	5	[ad, ad, ad]
RY	5	1	[ng, ng, ng, db, ng]
SG	3	1	[oo, ab, ab]
SM	3	2	[vi, ep, ou]
SY	3	1	[og, eb, nk]
TA	3	5	[pi, pi, pi]
VB	8	1	[he, dw, he, dw, he, he, he, he]
VW	3	10	[re, re, re]
WZ	3	1	[sh, ic, sh]
XA	3	3	[ti, ti, aw]
XM	3	3	[fh, tu, sy]
XR	4	1	[an, an, fm, an]
XW	4	1	[si, te, as, as]
XZ	5	1	[th, th, th, sl, th]
ZJ	5	5	[he, he, he, he, he]

A	0	0	8	1	9
B	5	10	15	0	0
C	0	9	2	1	1
D	3	0	2	0	0
E	0	0	5	4	0
F	5	6	10	1	5
G	4	4	0	3	8
H	3	11	1	3	3
I	2	3	3	15	0
J	11	2	0	2	5
K	1	3	5	1	7
L	3	0	5	3	5
M	5	3	2	7	3
N	11	0	0	0	0
O	0	9	2	1	4
P	1	3	7	0	0
Q	4	1	8	1	2
R	2	4	2	11	0
S	4	5	2	8	4
T	6	0	1	3	1
U	2	2	0	0	4
V	0	8	0	9	4
W	4	3	0	5	10
X	9	0	9	12	2
Y	4	2	2	1	5
Z	4	5	2	1	10

letter	$\equiv 1$ freq	$\equiv 2$ freq	$\equiv 3$ freq	$\equiv 4$ freq	$\equiv 0$ freq
A	0	0	8	1	9
B	5	10	15	0	0
C	0	9	2	1	1
D	3	0	2	0	0
E	0	0	5	4	0
F	5	6	10	1	5
G	4	4	0	3	8
H	3	11	1	3	3
I	2	3	3	15	0
J	11	2	0	2	5
K	1	3	5	1	7
L	3	0	5	3	5
M	5	3	2	7	3
N	11	0	0	0	0
O	0	9	2	1	4
P	1	3	7	0	0
Q	4	1	8	1	2
R	2	4	2	11	0
S	4	5	2	8	4
T	6	0	1	3	1
U	2	2	0	0	4
V	0	8	0	9	4
W	4	3	0	5	10
X	9	0	9	12	2
Y	4	2	2	1	5
Z	4	5	2	1	10

Table 3: Frequencies of the various Caesar shifts

letter	freq	freq	freq	freq	freq	total
	$\equiv 1 (5)$	$\equiv 2 (5)$	$\equiv 3 (5)$	$\equiv 4 (5)$	$\equiv 0 (5)$	
A	0	0	8	1	9	18
B	5	10	15	0	0	30
C	0	9	2	1	1	13
D	3	0	2	0	0	5
E	0	0	5	4	0	9
F	5	6	10	1	5	27
G	4	4	0	3	8	19
H	3	11	1	3	3	21
I	2	3	3	15	0	23
J	11	2	0	2	5	20
K	1	3	5	1	7	17
L	3	0	5	3	5	16
M	5	3	2	7	3	20
N	11	0	0	0	0	11
O	0	9	2	1	4	16
P	1	3	7	0	0	11
Q	4	1	8	1	2	16
R	2	4	2	11	0	19
S	4	5	2	8	4	23
T	6	0	1	3	1	11
U	2	2	0	0	4	8
V	0	8	0	9	4	21
W	4	3	0	5	10	22
X	9	0	9	12	2	32
Y	4	2	2	1	5	14
Z	4	5	2	1	10	22

Table 4: Frequencies of the various Caesar shifts

A	11000
B	10011
C	01110
D	10010
E	10000
F	10110
G	01011
H	00101
I	01100
J	11010
K	11110
L	01001
M	00111
N	00110
O	00011
P	01101
Q	11101
R	01010
S	10100
T	00001
U	11100
V	01111
W	11001
X	10111
Y	10101
Z	10001
Letters	11111
Figures	11011
Line feed	01000
Carriage return	00010
Word space	00100
All space	00000

Table 5: International teleprinter code