Exercise sheet 1, 16 November 2023

There are several nice tools online for cryptanalysis of classical systems, e.g. http://www.cryptool-online.org

There are more challenges online at https://www.mysterytwisterc3.org/en/

The raw data for the texts is also at http://www.hyperelliptic.org/tanja/teaching/CS15/data1.html. See below for some frequency distribution.

1. The Caesar cipher encrypts by shifting each letter by a fixed distance in the alphabet, where the letters are arranged so that Z is followed by A. The secret key that the sender and receiver share is the shifting distance (whether A maps to D or A maps to N or ....). The frequency distribution of text in English has clearly visible peaks at E, T, A, O, I, and N, and relatively high adjacent peaks at R, S, and T (see below for statistics) and encrypting with the Caesar cipher just shifts these peaks, so for long texts the shifting distance is best found by using statistics on the distribution of letters and matching that up with the natural distribution.

   Other common tools are to look for digraphs and trigraphs – combinations of two or three letters that appear frequently together in natural language, such as TH or even THE in THE, THEY, THERE, THEM, THEIR, OTHER, ...

   With computer support it is of course also possible to just try all 26 shifts and see which one gives readable text.

   (a) Find the plaintext for the following text was encrypted using the Caesar cipher.

   ```
   aopza leapz huleh twslv muvyt hsale azvao lbzbh skpza ypiba pvuzv
   mjohy hjaly zmvyl unspz oalea zovbs kovsk
   ```

   (b) Find the plaintext for the following text was encrypted using the Caesar cipher.

   ```
   drovo ddoba nyocx ydkzz okbyp doxex voccs xaekb dobae kbkxd sxoae
   oloma ekbdj ybaek cskbd spsms kvaeo cdsyx c
   ```

2. The Playfair cipher uses a keyword. Encryption/decryption uses a 5×5 grid of letters. To turn the keyword into this grid, start filling in the letters of the keyword row-wise from the top left corner. The grid contains each letter once, with I and J identified; so when you reach a letter in the keyword that has been used already, you skip it.

   After the end of the keyword, the remaining letters of the alphabet are inserted, again in the order they appear.

   If the keyword is SECRET then the grid looks as follows:

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

Note the skipped second E in the keyword. This fills up the grid completely – if you have any letters left, something went wrong earlier.

To encrypt a message, split it up into pairs of letters, starting from the beginning. If one pair consists of two equal letters, insert an X; this shifts the second letter into the next pair. If there is a single character at the end, append an X. There are three cases for encrypting a pair of letters:

(a) If the two letters appear in the same row, encrypt each of the two letters to the letter to the right of it.

(b) If the two letters appear in the same column, encrypt each of the two letters to the letter below it.

(c) If the two letters span a rectangle in the grid, encrypt each of them to the letter in the same row and opposite corner.

To decrypt, reverse the procedure.

The following text was encrypted using the Playfair cipher with keyword MATHEMATICS. Decrypt the message.

```
gc lz po nt au tc ad uh st
```

3. The Hill cipher is a secret-key system based on matrices. It takes a message in the English alphabet (26 characters), translates the characters into numbers as given below, and then encrypts the message by encrypting n numbers at a time as follows:
Let the secret key M be an $n \times n$ matrix over $\mathbb{Z}/26\mathbb{Z}$ which is invertible and let the plaintext a be the vector $(a_1, a_2, \ldots, a_n) \in (\mathbb{Z}/26\mathbb{Z})^n$. The corresponding ciphertext is $c^T = Ma^T$. To decrypt compute $a^T = M^{-1}c^T$.
Reminder: not every element in $\mathbb{Z}/26\mathbb{Z}$ is invertibe. A matrix is invertible if its determinant is invertible.

(a) Let

$$M = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 3 & 1 \end{pmatrix}.$$

Encrypt the text `CRY PTO`

(b) Let M be a 2×2 matrix. You know that $(1, 3)^T$ was encrypted as $(-9, -2)^T$ and that $(7, 2)^T$ was encrypted as $(-2, 9)^T$. Find the secret key M.

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

4. The affine encryption system is a symmetric system. The key consists of two integers $0 \leq a, b < 31$. Messages and ciphertexts are also integers in $[0, 30]$. Message m is encrypted as $c = a \cdot m + b \bmod 31$.

   You observe two ciphertexts for which you know the plaintexts:

   - $m_0 = 17$ encrypts to $c_0 = 23$.
   - $m_1 = 5$ encrypts to $c_1 = 26$

   Find the key and decrypt the ciphertext $c_2 = 1$.

Probability distributions of 1-grams in English, from Henk van Tilborg "Fundamentals of cryptology", page 5. Boldfacing of values larger than 0.06 by me. Note the probabilities of e and the triple r,s, and t.

a	**0.0804**	b	0.0154	c	0.0306	d	0.0399
e	**0.1251**	f	0.0230	g	0.0196	h	0.0549
i	**0.0726**	j	0.0016	k	0.0067	l	0.0414
m	0.0253	n	**0.0709**	o	**0.0760**	p	0.0200
q	0.0011	r	**0.0612**	s	**0.0654**	t	**0.0925**
u	0.0271	v	0.0099	w	0.0192	x	0.0019
y	0.0173	z	0.0009				