Codebusters @ UChicago's SO-ED Fall 2018

October 13

Lisa Lin and Oliver Tsang

Lesson Plan:

- Overview of cryptography
  - Don't want other people to understand what you say
  - Want your friend to understand what you say
  - Many ciphers exist -- each with different tradeoffs
  - Turns out a lot of systems can be broken even without knowledge of the keys
- See several examples of ciphers in use
  - Caesar cipher: encoding, decoding, and breaking the code
  - Affine cipher: encoding, decoding (with special case of Atbash cipher)
  - Vigenère cipher: encoding and decoding
- Tips for practicing
  - Memorize the numbers of each alphabet letter as well as approximate frequencies in English (and eventually Spanish)
  - Don't worry about time at first, but try to speed up when you get comfortable.
  - Use the computer (esp. online tools, e.g. practicalcryptography.com) to generate ciphertext faster!

Frequently used information about the alphabet (frequency is for English)

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M
Number	1	2	3	4	5	6	7	8	9	10	11	12	13
Frequency (%)	8.1	1.5	2.7	4.3	12.0	2.3	2.0	5.9	7.3	0.1	0.7	4.0	2.6

Letter	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Number	14	15	16	17	18	19	20	21	22	23	24	25	26
Frequency (%)	7.0	7.7	1.8	0.1	6.0	6.3	9.1	2.9	1.1	2.1	0.2	2.1	0.1

Example Problems

Part 1: Caesar Cipher

To encode or decode a message using the Caesar cipher, replace each letter by the letter that comes a fixed number of places later in the alphabet. Classically, a shift of 3 was used so that A becomes D, B becomes E, ..., X becomes A, Y becomes B, and Z becomes C.

1.1) Try encoding the following message with a shift of 15 (or -11 if you prefer):

Do you want me to send you back to where you were, unemployed in Greenland?

1.2) Decode the following message, which was encoded with a shift of 5:

SJAJW LT NS FLFNSXY F XNHNQNFS BMJS IJFYM NX TS YMJ QNSJ!

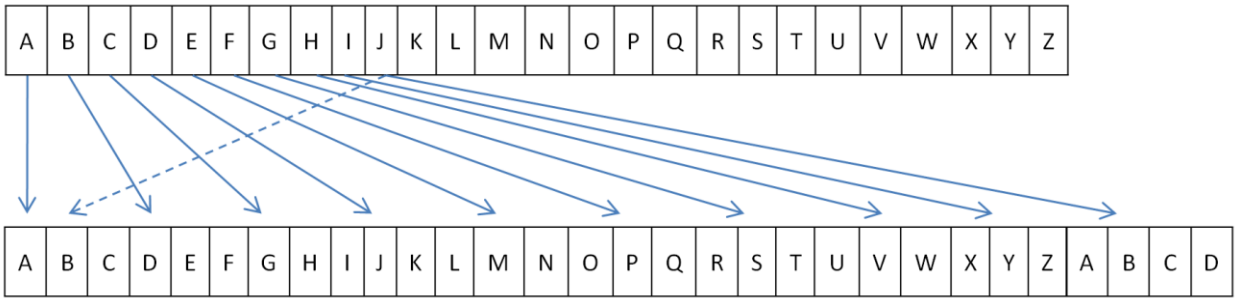
1.3) Break the following code with the help of the frequency table below:

AOLYL DLYL THUF LUPNTH THJOPULZ BZLK KBYPUN DVYSK DHY PP, HUK AOLF DLYL JVUZAHU ASF TVKPMPLK IF AOL NLYTHUZ.

Letter	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
Frequency	5	2	0	4	0	4	0	7	1	2	6	13	1	3	4	7	0	0	2	5	8	3	0	0	7	4

Part 2: Affine Cipher

As with the Caesar cipher, we shift letters over. However, we can skip several letters between each letter. Below is an illustration of what happens to the first few letters when we have no offset and a skip of 3:



This can be described more formally by representing each letter by a number/variable. For affine ciphers, usually we assign 0 to 'a,' 1 to 'b,' etc. rather than starting at 1 as before. Below, E(x) represents encryption, while D(y) represents decryption.

$$E(x) = ((a \cdot x + b) \bmod 26)$$

$$D(y) = \left(a^{-1} \cdot (y - b) \bmod 26\right)$$

You can check that D(E(x)) will return the original x.

Mathematicians like to call things in the form ax+b "affine." The mod 26 just means that we wrap around the end of the alphabet if necessary. a^{-1} is a specially chosen number ("multiplicative inverse") such that $a \cdot a^{-1} \equiv 1 \pmod{26}$. For this number to

exist, *a* must be relatively prime with 26. That is, *a* cannot be a multiple of 2 or 13 (26's factors). Here is a table of multiplicative inverses modulo 26 for your convenience:

a	1	3	5	7	9	11	13	15	17	19	21	23	25
a^{-1}	1	9	21	15	3	19	N/A	7	23	11	5	17	25

2.1) Encode "*green eggs and ham*" with $a = 3$ and $b = 19$ (equivalently, $b = -7$).

2.2) Decode the following message using $a = 7$ and $b = 23$:

URKZTASBTXNMZZDZKTBTRGAZKXNRRSTB

NKRGYMRNMZTT.

2.3) The Atbash cipher is just a simple case of the Affine cipher, where $a = b = 25$ (or -1). Try this:

Z NZM Z KOZM Z XZMZO KZMZNZ.

Part 3: Vigenère Cipher (roughly pronounced “Veej-nair”)

The Vigenère cipher effectively has a different caesar shift for each letter. These shifts are typically given by a keyword shorter than the plaintext message, so you repeat the code word. Here’s an example of encoding “Codebusters” using the keyword “scioly” to give the offsets:

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

As you can see, the final output becomes *uqlsmskvmfd*. Despite the fact that the letters ‘e’ and ‘s’ appear twice, they are mapped onto different letters each time. This makes the Vigenère cipher more difficult to break by frequency analysis. (Note: if you know the length of the key word, you

can break the message into different portions and perform frequency analysis to break each letter of the keyword independently).

3.1) Encode the following message using the keyword *astro*:

Yury says, “Yay, syzygy!”

3.2) Decode the following message using the keyword *monty:*

RVUVZPDBMONBOHCFPIMZTT