Intermediate Math Circles
March 19, 2014
Cryptography I

Introduction to Cryptography

Cryptography = the study of sending messages in secret form
Cryptanalysis = the study of breaking cryptographic messages
Cryptology = the above two things together.

The need for secret communications has been known for centuries:

- **Scytale** (a **Transposition Cipher**) dates to 5th century BC, used by Spartan military

- **Caesar's Shift Cipher** was used by Julius Caesar in 50 B.C. to communicate with his military

Just as there is a need for being able to send secret messages, there is also a need to read secret messages.

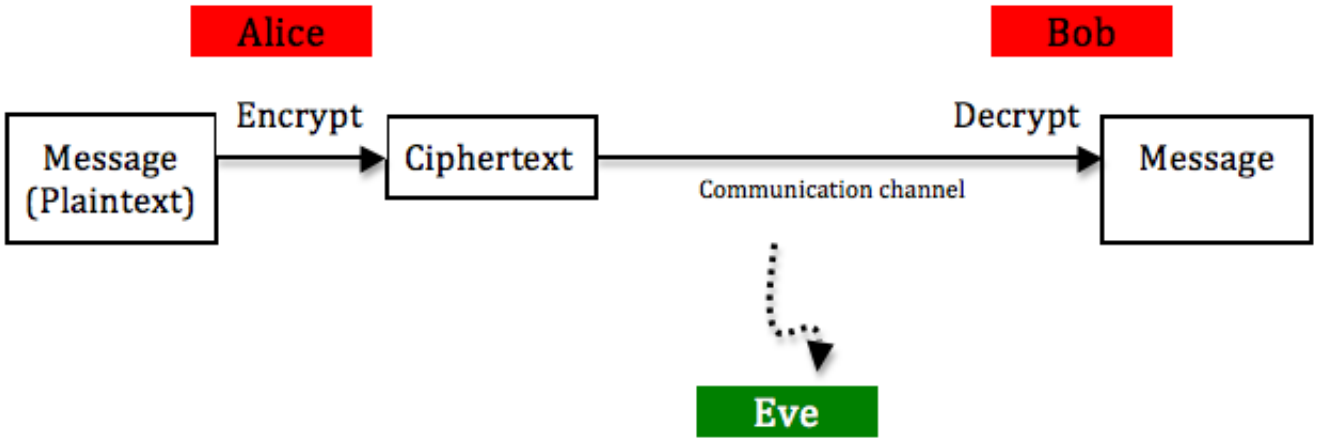
In the 20th century, cryptography played a significant role in many global conflicts (eg. Enigma machine and Bletchley Park in WWII).

There was a historical need for cryptography in military and government areas and there still is, but in the modern world the need for secret communication is much larger than it was even in the recent past.

The credit card, debit card and web transactions of modern commerce, as well as privacy concerns for the electronic storage of health, citizenship and other records, have raised the need for secure communications and secure storage dramatically.

Setup:

We have a sender, Alice, and a receiver, Bob.
Alice wants to send a message M to Bob, over an insecure channel, but she wants only Bob to be able to read the message.
There is a good chance that at least part of the transmitted message will be intercepted by and eavesdropper, Eve.
Alice needs to "encrypt" the message so that Eve cannot read it.
Alice wants to use a simple algorithm, so that Bob can "decrypt" the transmitted message using some special key that only he has, but so that it is hard for Eve to "break the code" without knowing the key.



Plaintext = original message
Ciphertext = encrypted message
Encryption = act of transforming plaintext into ciphertext
Encryption algorithm = method used to turn plaintext into ciphertext. Uses a **key**, some input into the algorithm.
Decryption = act of transforming cipher text into plaintext
Decryption algorithm = method used to turn ciphertext into plaintext. Uses a **key**, some input into the algorithm.

The security of the code should depend on keeping the **key** a secret, not keeping the encryption algorithm a secret.

Alice should be able to code her message with a well-known method and still be reasonably confident that the message cannot be decoded by anyone other than Bob, since he is the only person with the key.

First some Modular Arithmetic

Question: What is the remainder when 51 is divided by 7? (Answer: 2).

We write $51 \equiv 2 \mod 7$

This means that 51 and 2 have the same remainder when divided by 7.
In other words, 51 and 2 differ by a multiple of 7, or $51 - 2 = 49$ is a multiple of 7.

In general, if a, b and m are integers, $a \equiv b \mod m$, "a is **congruent** to b modulo m" means that a and b differ by a multiple of m, or $a - b = km$, where k is some integer.

We will be interested in the smallest integer $b \geq 0$ such that $a - b = km$, where k is some integer.

Example 1: Reduce

a) $52 \mod 8$
b) $41 \mod 5$
c) $84 \mod 4$
d) $-17 \mod 4$
e) $145672 \mod 13$

Modular Arithmetic:

$28 \equiv 4 \pmod 8$ and $11 \equiv 3 \pmod 8$.
What happens when we add or subtract?

$28 + 11 = 39 \equiv 7 \pmod 8$.
Notice: $3 + 4 = 7 \equiv 7 \mod 8$

$28 - 11 = 17 \equiv 1 \pmod 8$.
Notice: $4 - 3 = 1 \equiv 1 \mod 8$

In general: If $a \equiv A \pmod m$ and $b \equiv B \pmod m$, then

(i) $a + b \equiv A + B \pmod m$
(ii) $a - b \equiv A - B \pmod m$

Example 2: Reduce (in two different ways)

a) $17 + 21 \mod 6$
b) $83 - 21 \mod 3$
c) $21 - 83 \mod 11$

Exercise Set 1:

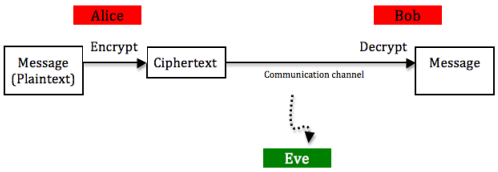
1. Reduce 237288 modulo 5
2. Reduce $192 + 118$ modulo 5
3. Reduce $192 - 118$ modulo 5
4. Reduce $118 - 192$ modulo 5
5. Today is a Wednesday. What day of the week will it be
   a) 100 days from now?
   b) 365 days from now?
   c) 1000 days from now?
6. Emily celebrated her 13th birthday on Wednesday, February 19th, 2014. On what day of the week was she born? (Don't forget about the leap years in 2004, 2008 and 2012!)

Answers to Exercise Set 1:

1. $237288 \equiv 3 \mod 5$
2. $192 + 118 \equiv 0 \mod 5$
3. $192 - 118 \equiv 4 \mod 5$
4. $118 - 192 \equiv 1 \mod 5$
5. a) Friday
   b) Thursday
   c) Tuesday
6. Monday

Substitution Ciphers: A couple of examples

Recall: We have a sender, Alice, and a receiver, Bob.
Alice wants to send a message M to Bob, so that only Bob can decipher the message.
However, there is a good chance that at least part of the transmitted message will be intercepted by and eavesdropper, Eve.
So, Alice needs to "encrypt" the message so that Eve cannot read it.



The Caesar Shift Cipher:

Assign the numbers 0 to 25 to the letters A to Z (so A is 0, B is 1 and so on, Z is 25).
Think of the alphabet mod 26.
Pick a random number to be your **key**, call it k.

Encryption Algorithm: Encrypt each letter individually using the formula:

$$\text{coded} = (\text{original} + k) \pmod{26}$$

Decryption Algorithm: Decrypt each letter individually using the formula:

$$\text{original} = (\text{coded} - k) \pmod{26}$$

Suppose we pick $k = 7$.
A is encrypted as H since $0 + 7 \equiv 7 \mod 26$
B is encrypted as I since $1 + 7 \equiv 8 \mod 26$
H is encrypted as O since $7 + 7 \equiv 14 \mod 26$
U is encrypted as B since $20 + 7 = 27 \equiv 1 \mod 26$
and so on. We get the following permutation of the alphabet:

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
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

So plaintext CRYPTOFUN is encrypted as JYFWAVMBU.

How does Bob decrypt the ciphertext into plaintext?
To decrypt this message, Bob just needs to know the secret key.

Example 3:

a) Using a Caesar Shift Cipher and secret key 11, encrypt the message "I WANT COOKIES". (Answer: T HLYE NZZVTPD)

b) Using a Caesar Shift Cipher and secret key 11, decrypt the message "NSPNV ESP NFAMZLCO". (Answer: Check the cupboard)

Exercise Set 2:

1. Encode the message "MODULAR ARITHMETIC" using a Caesar Shift cipher with secret key $k = 7$.
2. Decode the message "SLA AOLT LHA JHRL" using a Caesar Shift cipher with secret key $k = 7$.
3. Encode the message "ALL SQUARES ARE RECTANGLES" using a Caesar Shift cipher with secret key $k = 14$.
4. Decode the message "PIH BCH OZZ FSQHOBUZSG OFS GEIOFSG" using a a Caesar Shift cipher with secret key $k = 14$.
5. Find a partner to work with. Think of a secret message to send to them and encode it using a Caesar Shift cipher with a secret key of your choice. Give your partner the coded message and shift number. Decode your partner's message to you.

Answers to Exercise Set 2:

1. TVKBSHY HYPAOTLAPJ
2. LET THEM EAT CAKE
3. OZZ GEIOFSG OFS FSQHOBUZSG
4. BUT NOT ALL RECTANGLES ARE SQUARES

Encryption and Decryption are very easy with the Caesar Shift Cipher. Sadly, this it is also very easy to break. Can you see how?

Just to try all possible keys (there are only 26 - don't even need a computer for this!).

How can we improve on the Caesar shift cipher?

Instead of shifting the alphabet, use a random permutation of the alphabet to get a **Substitution Cipher**.

For example:

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

The table above acts as the **key** for this cipher.

How many possible keys are there? (Answer: $26! \approx 4 \times 10^{26}$)
Breaking the cipher by trying all keys is no longer feasible, even for a computer!

Example 4:

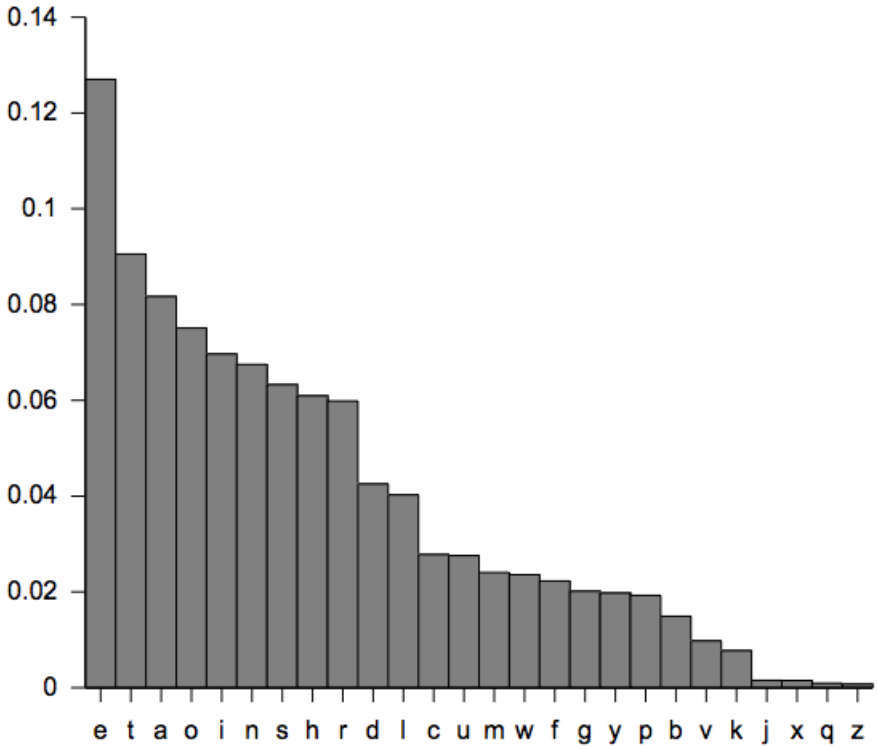
a) Using the Substitution Cipher above, encrypt the message "I WANT COOKIES".
(Answer: B LOXD CEEUBKT)

b) Using the Substitution Cipher above, decrypt the message "BX DWK CEEUBK QOH".
(Answer: In the cookie jar)

Breaking a Substitution Cipher:

The Substitution Cipher is better than the Caesar Shift Cipher, but unfortunately, it can also be easily broken.

How?



A statistical analysis can be used, using known letter frequencies in the English alphabet.

Order of relatively frequency:

1. E
2. T, A, O, I, N, S, H, R
3. D, L
4. C, U, M, W, F, G, Y P, B
5. V, K, J, X, Q, Z

Exercise Set 3:

Using known letter frequencies in the English alphabet, try to break the code below.
The message was encrypted with a substitution cipher.

```
IFYYOL PYZZR AXRGVK QWBZ IQL IFWK FB ZWEV PXB QB MFL ZWV ZJ BIV JVM
BIFB KQK. IVYRQZWV DYFWDVYL IFK LQRGUO YZUUVK ZHVY ZW BIV DYZXWK
FWK WVHQUUVL IFKWB RZHVK FB FUU. GVYIFGL PYZZRL UQTV IZYLVL EZXUK
BVUU MIVW OZX MVYV FJYFQK BIZXDIB IFYYO BIVYV MFL F SXFHVY QW
WVHQUUVL HZQEV BIFB LFQK ZWUO BZZ EUVFYUO BIFB IV MFWBVK BZ TVVG
IQL JVVB ZW BIV DYZXWK.
RFKFR IZZEI BIVW LIZMVK BIVR IZM BZ RZXWB BIVQY PYZZRL MQBIZXB
LUQKQWD ZJJ BIV VWK FWK MFUTVK XG FWK KZMW BIV YZML EZYYVEBQWD
BIVQY DYQGL.
```

Here are the frequencies of each letter in this example:

A	B	C	D	E	F	G	H	I	J	K	L	M
1	34	0	8	7	28	7	6	32	6	23	19	12

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	6	4	21	11	1	3	18	45	26	11	27	38

IFYYOL PYZZR AXRGVK QWBZ IQL IFWK FB ZWEV PXB QB MFL ZWV ZJ BIV JVM BIFB KQK. IVYRQZWV DYFWDVYL IFK LQRGUO YZUUVK ZHVY ZW BIV DYZXWK FWK WVHQUUVL IFKWB RZHVK FB FUU. GVYIFGL PYZZRL UQTV IZYLVL EZXUK BVUU MIVW OZX MVYV FJYFQK BIZXDIB IFYYO BIVYV MFL F SXFHVY QW WVHQUUVL HZQEV BIFB LFQK ZWUO BZZ EUVFYUO BIFB IV MFWBVK BZ TVVG IQL JVVB ZW BIV DYZXWK.

RFKFR IZZEI BIVW LIZMVK BIVR IZM BZ RZXWB BIVQY PYZZRL MQBIZXB LUQKQWD ZJJ BIV VWK FWK MFUTVK XG FWK KZMW BIV YZML EZYYVEBQWD BIVQY DYQGL.