COMP170 Discrete Mathematical Tools for Computer Science

Intro to Crypto and Mod

Version 2.0: Last updated, May 13, 2007

Discrete Math for Computer Science K. Bogart, C. Stein and R.L. Drysdale Section 2.1, pp. 43-54

Godfrey Harold Hardy

b. 1877. d. 1947

was a prominent British mathematician, known for his achievements in number theory and mathematical analysis. from http://en.wikipedia.org/wiki/G._H._Hardy

In his 1940 autobiography

A Mathematician's Apology, Hardy wrote



"The great modern achievments of applied mathematics have been in relativity and quantum mechanics, and these subjects are, at present, almost as 'useless' as the theory of numbers."

and

"... then the great bulk of higher mathematics is useless. Modern Geometry and algebra, the theory of numbers, the theory of aggregates and functions, relativity, quantum mechanics — no one of them stands the test much better than another, . . ."

If he could see the world now, G.H. Hardy would be spinning in his grave.

Number theory, introduced in this lecture, is the basis of modern coding theory. Computer security and e-commerce would be *impossible* without it.

relativity and quantum theory turned out to be pretty useful as well

At one point, not long ago, the largest employer of mathematicians in the United States, and therefore probably the world, was the National Security Agency (NSA). The NSA is the largest spy agency in the US – bigger than the CIA – and has the responsibility for code design and breaking.

(Euclid's Division Theorem)

Let n be a positive integer. Then for every integer m, there exist unique integers q and r such that m = nq + r and $0 \le r < n$.

This will be proven in next lecture. It says that $m \mod n$ is *uniquely* defined.

2.1 Cryptography and Modular Arithmetic

- Arithmetic Modulo n
- Introduction to Cryptography
- Private-Key Cryptography
 - ullet Caesar Ciphers: Cryptography Using Addition $\bmod n$
 - Cryptography Using Multiplication mod n
- Public-Key Cryptography

A quick review of the laws of arithmetic over the real numbers

• The *commutative laws* for addition and multiplication $a+b=b+a; \quad ab=ba$ Ex: $3+7.2=7.2+3; \quad 3\cdot 5=5\cdot 3.$

• The associative laws for addition and multiplication

$$a + (b + c) = (a + b) + c;$$
 $a(bc) = (ab)c$
Ex: $5 + (3 + 7) = (5 + 3) + 7;$ $5 \cdot (3 \cdot 7) = (5 \cdot 3) \cdot 7$

• The distributive law

$$(a+b)c = ac + bc$$

 $(5+3) \cdot 7 = 5 \cdot 7 + 3 \cdot 7$

- Every number a has an additive inverse -a such that a+(-a)=0. Ex: 5+(-5)=0.
- Every number $a \neq 0$ has a multiplicative inverse a^{-1} s.t. $aa^{-1} = 1$. Ex: $5 \cdot \frac{1}{5} = 1$.

2.1 Cryptography and Modular Arithmetic

- Arithmetic Modulo *n*
- Introduction to Cryptography
- Private-Key Cryptography
 - ullet Caesar Ciphers: Cryptography Using Addition $\bmod n$
 - Cryptography Using Multiplication mod n
- Public-Key Cryptography

Modular Arithmetic

Definition (1st version);

Let m, n be positive integers. Then $m \mod n$ is the remainder left when dividing m by n.

Ex: $25 \mod 4 = 1$; $25 \mod 5 = 0$; $27 \mod 5 = 2$.

We would like to extend this definition to negative m as well.

Definition (2nd version);

For an integer m and positive integer n, $m \bmod n$ is the smallest nonnegative integer r such that, for some integer q, m = nq + r.

Ex: $25 \mod 4 = 1$; $-25 \mod 4 = 3$.

Definition (2nd version); For an integer m and positive integer n, $m \bmod n$ is the smallest nonnegative integer r such that, for some integer q, m = nq + r.

 $25 \mod 4 = 1$ because $25 = 4 \cdot 6 + 1$ and any other way of writing $25 = 4 \cdot q + r$ would have an r bigger than 1.

 $-25 \mod 4 = 3$ because $-25 = 4 \cdot (-7) + 3$ and any other way of writing $-25 = 4 \cdot q + r$ would have an r bigger than 3. (why?)

Note: In general, except if
$$[m \mod n] = 0$$
,
$$[(-m) \mod n] = n - [m \mod n] \quad \text{so}$$

$$[(-m) \mod n] \neq [m \mod n] \quad \text{unless}$$

$$[m \mod n] = n/2.$$

Arithmetic Modulo n

```
Compute
21 \mod 9
38 \mod 9
 (21 \cdot 38) \mod 9
 (21 \bmod 9) \cdot (38 \bmod 9)
 (21 + 38) \mod 9
 (21 \bmod 9) + (38 \bmod 9)
It looks as if \lceil (ab) \bmod n \rceil = \lceil (a \bmod n) \cdot (b \bmod n) \rceil
       and [(a + b) \mod n] = [(a \mod n) + (b \mod n)]
Is this true? No! Try a = 2, b = 8, n = 9.
```

So what is happening here?

True or false?

$$i \bmod n = (i+2n) \bmod n$$
? $i \bmod n = (i-3n) \bmod n$?

Both true, since adding multiples of n to i does not change the value of the *remainder*, $i \mod n$.

Lemma 2.2

 $i \bmod n = (i + kn) \bmod n$ for all integers k.

Proof:

- By Euclid's Division Theorem, i = nq + r (*), for *unique* integers q and r, with $0 \le r < n$.
- By (*) and definition of mod, $r = i \mod n$.
- Adding kn to both sides, i + kn = n(q + k) + r (**).
- From (**), Eucid's div thm and definition of mod, $r = (i + kn) \mod n$, and we are done.

Lemma 2.3

```
(i+j) \mod n = (i+(j \mod n)) \mod n= ((i \mod n) + j) \mod n= ((i \mod n) + (j \mod n)) \mod n,
```

$$(i \cdot j) \mod n = (i \cdot (j \mod n)) \mod n$$

= $((i \mod n) \cdot j) \mod n$
= $((i \mod n) \cdot (j \mod n)) \mod n$.

Lemma 2.3

$$(i+j) \mod n = (i+(j \mod n)) \mod n$$
$$= ((i \mod n) + j) \mod n$$
$$= ((i \mod n) + (j \mod n)) \mod n,$$

Proof:

We prove that item on left is equal to bottom item on right. Proofs of all other equalities are very similar.

By Euclid's Division Theorem, for unique q_1 and q_2 , $i = (i \mod n) + q_1 n$ and $j = (j \mod n) + q_2 n$.

Adding these 2 equations together mod n and using Lemma 2.2,

$$(i + j) \mod n = ((i \mod n) + q_1 n + (j \mod n) + q_2 n) \mod n$$

= $((i \mod n) + (j \mod n) + n(q_1 + q_2)) \mod n$
= $((i \mod n) + (j \mod n)) \mod n$.

Definition:

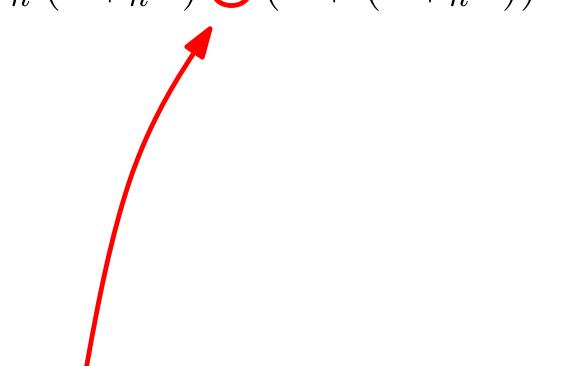
 Z_n is the set of integers $\{0,1,\ldots,n-1\}$ with addition $\operatorname{mod} n$ $i+_n j=(i+j) \operatorname{mod} n$ and multiplication $\operatorname{mod} n$ $i\cdot_n j=(i\cdot j) \operatorname{mod} n$

- If $x \in \mathbb{Z}_n$, then x is a variable with possible integral values between 0 and n-1.
- If $x, y \in Z_n$, we use $x +_n y$ and $x \cdot_n y$ to perform algebraic operations on x, y.
- Additive identity property: $0 +_n i = i$. Multiplicative identity property: $1 \cdot_n i = i$.
- $a -_n b$ denotes $a +_n (-b)$.

Addition and multiplication mod n satisfy the **commutative**, **associative** and **distributive** laws.

Proof: Commutativity of $+_n$ and \cdot_n follows immediately from commutativity of ordinary addition and multiplication. We prove the associative law for addition in the following equations; the other laws follow similarly.

$$a +_n (b +_n c) \bigoplus (a + (b +_n c)) \mod n$$



$$i +_n j = (i + j) \mod n$$
 and $i \cdot_n j = (i \cdot j) \mod n$.

$$a +_n (b +_n c) = (a + (b +_n c)) \mod n$$

$$\Rightarrow (a + ((b + c) \mod n)) \mod n$$

$$i +_n j = (i + j) \mod n$$
 and $i \cdot_n j = (i \cdot j) \mod n$.

$$a +_n (b +_n c) = (a + (b +_n c)) \mod n$$
$$= (a + (b + c) \mod n)) \mod n$$
$$\bigoplus (a + (b + c)) \mod n$$



Addition and multiplication mod n satisfy the **commutative**, **associative** and **distributive** laws.

$$a +_n (b +_n c) = (a + (b +_n c)) \mod n$$

$$= (a + (b + c) \mod n)) \mod n$$

$$= (a + (b + c)) \mod n$$

$$\bigoplus ((a + b) + c) \mod n$$

Associative law for ordinary sums.

Addition and multiplication mod n satisfy the **commutative**, **associative** and **distributive** laws.

$$a +_n (b +_n c) = (a + (b +_n c)) \mod n$$

$$= (a + ((b + c) \mod n)) \mod n$$

$$= (a + (b + c)) \mod n$$

$$= ((a + b) + c) \mod n$$

$$\bigoplus ((a + b) \mod n + c) \mod n$$

Lemma 2.3

$$a +_n (b +_n c) = (a + (b +_n c)) \mod n$$

$$= (a + (b + c) \mod n)) \mod n$$

$$= (a + (b + c)) \mod n$$

$$= ((a + b) + c) \mod n$$

$$= ((a + b) \mod n + c) \mod n$$

$$\bigoplus ((a +_n b) + c) \mod n$$

$$i +_n j = (i + j) \mod n$$
 and $i \cdot_n j = (i \cdot j) \mod n$.

Addition and multiplication mod n satisfy the **commutative**, **associative** and **distributive** laws.

$$a +_n (b +_n c) = (a + (b +_n c)) \mod n$$

$$= (a + (b + c) \mod n)) \mod n$$

$$= (a + (b + c)) \mod n$$

$$= ((a + b) + c) \mod n$$

$$= ((a + b) \mod n + c) \mod n$$

$$= ((a +_n b) + c) \mod n$$

$$= (a +_n b) +_n c$$

 $i +_n j = (i + j) \mod n$ and $i \cdot_n j = (i \cdot j) \mod n$.

2.1 Cryptography and Modular Arithmetic

- Arithmetic Modulo n
- Introduction to Cryptography
- Private-Key Cryptography
 - ullet Caesar Ciphers: Cryptography Using Addition $\bmod n$
 - Cryptography Using Multiplication mod n
- Public-Key Cryptography

Introduction to Cryptography

Cryptography is the study of methods for sending and receiving secret messages.

A sender is trying to send a message to a receiver. An adversary wants to steal the message.

Sensitive information is **encrypted** — modified in such a way that it should be only understandable by the receiver and **undecipherable** to the adversary.

The method is deemed **successful** if the sender is able to communicate a messsage to the receiver without the adversary learning what that message was.

A difficult goal!

2.1 Cryptography and Modular Arithmetic

- Arithmetic Modulo n
- Introduction to Cryptography
- Private-Key Cryptography
 - ullet Caesar Ciphers: Cryptography Using Addition $\bmod n$
 - Cryptography Using Multiplication mod n
- Public-Key Cryptography

Private-Key Cryptography

Sender and receiver agree in advance on a secret code and then send messages using that code.

Caesar cipher: A private-key cryptosystem in which letters of the alphabet are shifted (circularly) by some fixed amount.

This cipher is named after the Roman emperor **Julius Caesar** (b. 100BC, d. 44BC). Caesar supposedly used this type of cipher (with a shift of 3) to communicate with his generals.

Private-Key Cryptography

Sender and receiver agree in advance on a secret code and then send messages using that code.

Caesar cipher: A private-key cryptosystem in which letters of the alphabet are shifted (circularly) by some fixed amount.

Original message is called plaintext and the encoded text is called ciphertext.

```
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
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
Plaintext message: ONE IF BY LAND AND TWO IF BY SEA.

Ciphertext: SRI MJ FC PERH ERH XAS MJ FC WIE.
```

Easy to implement using arithmetic mod 26.

Use 0 for A, 1 for B, \dots Convert a message to a sequence of numbers. For example, SEA becomes $18\ 4\ 0$.

Q: What does the numerical representation of this word become if we shift every letter 2 places to the right?

A: Replace n by $(n+2) \mod 26$, \Rightarrow SEA becomes $20 \ 6 \ 2$.

Q: What if we shift every letter 13 places to the right?

A: Replace n by $(n+13) \mod 26$, \Rightarrow SEA becomes 5 17 13.

A Caesar cipher with shift s can easily be implemented on most computers by replacing each "letter" n with $(n+s) \mod 26$. Most computer languages can easily convert between text and numbers, and provide predefined \mod functions.

A Caesar cipher with shift s can easily be implemented on most computers by replacing each "letter" n with $(n+s) \mod 26$.

• E.G. To shift each letter 2 to the right, replace n by $n' = (n+2) \bmod 26$, \Rightarrow SEA becomes $20 \ 6 \ 2$.

How should the receiver decipher (decrypt) the message?

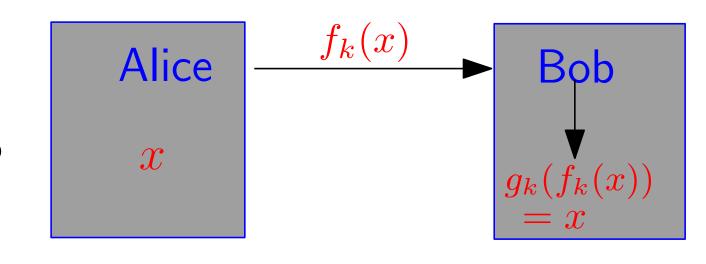
Easy: Replace n' by $(n'-s) \mod 26$,

- E.G. If s = 2, then a received 20 6 2 becomes $18 \ 4 \ 0$ which is SEA.
- So, we've just seen how $+_n$ on Z_n (for n=26) can be used to implement **encrypting** and **decrypting** Caesar ciphers.

A slightly different view

- A Caesar cipher has a private-key k
- To encode x, use the function $f_k(x) = x +_{26} k$
- To decode y, use the function $g_k(y) = y 26 k$
- Note that $g_k(y) = f_k^{-1}(y)$, i.e., g(f(x)) = x and f(g(y)) = y
- 0) Bob & Alice know k
- i) Alice has letter x
- ii) She sends $f_k(x)$ to Bob
- iii) Bob calculates

$$x = g_k(f_k(x))$$



2.1 Cryptography and Modular Arithmetic

- Arithmetic Modulo n
- Introduction to Cryptography
- Private-Key Cryptography
 - ullet Caesar Ciphers: Cryptography Using Addition $\bmod n$
 - ullet Cryptography Using Multiplication $\bmod n$
- Public-Key Cryptography

Cryptography Using Multiplication mod n

- We just saw how to use modular addition/subtraction to encrypt/decrypt. Now we'll discuss how to use modular multiplication. We assume that message is a *number*, M.
- Private key is some a, n.

Encrypt: $f_{a,n}(M) = a \cdot M \mod n = a \cdot_n M$

Decrypt: "Divide" f(M) by $a \mod n$.

• In order for this to work *division* must exist and define an inverse to multiplication., i.e.

$$g_{a,n}(f_{a,n}(M)) = M$$
 where $g_{a,n}(X) = a^{-1} \cdot_n X$

Does division exist?

What exactly does division mod n mean?

Suppose, for some x, we had calculated $f(x) = a \cdot_n x$. Does f^{-1} exist?

Consider the following three cases of a, x, n.

- (a) (a, x, n) = (4, 3, 12),
- (b) (a, x, n) = (3, 6, 12),
- (c) (a, x, n) = (5, 7, 12)

$$f(x) = a \cdot_n x$$

(a)
$$(a, x, n) = (4, 3, 12)$$
:

You send the message

$$f(x) = 4 \cdot_{12} 3 = 0$$
.

Recipient receives 0.

Problem:

There are 4 values of x, (0, 3, 6, 9), s. t. $a \cdot_{12} x = 0$.

Recipient doesn't know which x you intended.

 $\Rightarrow f^{-1}$ doesn't exist! Impossible to decrypt!

$$f(x) = a \cdot_n x$$

(b)
$$(a, x, n) = (3, 6, 12)$$
:

You send the message

$$f(x) = 3 \cdot_{12} 6 = 6$$
.

Recipient receives 6.

Problem:

There are 3 values of x, (2, 6, 10), s. t. $a \cdot_{12} x = 6$.

Recipient doesn't know which x you intended.

 $\Rightarrow f^{-1}$ doesn't exist! Impossible to decrypt!

$$f(x) = a \cdot_n x$$

(c)
$$(a, x, n) = (5, 7, 12)$$
:

You send the message

$$f(x) = 5 \cdot_{12} 7 = 11.$$

Recipient receives 11.

In fact,

7 is unique solution to

$$5 \cdot_{12} x = 11!$$

Recipient does know which x you intended.

⇒ Recepient could decrypt this messsage!

What exactly does division mod n mean?

Suppose, for some x, we had calculated $f(x) = a \cdot_n x$. Does f^{-1} exist?

Consider the following three cases of a, x, n.

- (a) (a, x, n) = (4, 3, 12), (b) (a, x, n) = (3, 6, 12),
- (c) (a, x, n) = (5, 7, 12)
 - We just saw that in cases (a) and (b), there is an $x' \neq x$ s.t. f(x') = f(x) so recipient would not be able to decrypt message. This means that we can not use this f(x) as an encoding function
 - In case (c), given x, recipient could uniquely calculate x so f(x) might be a good encoding function.
 - f(x) can be used as an encoding function when f(x) has an inverse!

When does $f_{a,n}(x) = a \cdot_n x$ have an inverse?

 $f_{a,n}(x) = a \cdot_n x$ has an inverse if and only if a and n are relatively prime, i.e., they have no common factors greater than 1.

In the next lecture we will see what this means and how to use it to define divsion in \mathbb{Z}_n .

2.1 Cryptography and Modular Arithmetic

- Arithmetic Modulo n
- Introduction to Cryptography
- Private-Key Cryptography
 - ullet Caesar Ciphers: Cryptography Using Addition $\bmod n$
 - Cryptography Using Multiplication mod n
- Public-Key Cryptography

- The Caesar cipher is a very simple form of private-key encryption. The *key* is the shift.
- There are many other more sophisticated and secure types of private-key encryption, see, e.g., COMP364
- The security of **all** of them rely on the fact that the codebook is kept secret between the sender and receiver.
- What happens if codebook is lost or stolen?
 No good answer to this
 How can we distribute different codebooks
 to millions of customers?
 Problem in e-commerce where each customer
 would need own codebook
- Motivation for Public-Key Cryptography

Public-Key Cryptosystems

 In private-key cryptosystems the sender and receiver share a private-key or codebook.

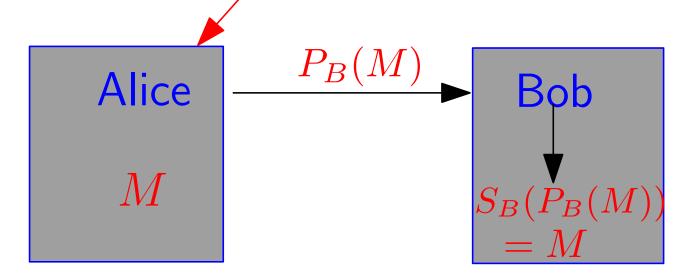
The same key is used for encypting and decrypting. Implicit assumption: knowing how a message is encypted implies knowing how to decrypt it

- In public-key cryptography this is no longer true. Everybody has two keys; a public key and a secret key.
- My public key: Known by all. Used to send me a message My secret key: Known only by me.
 Used to decrypt messages sent to me that were encrypted using my public key.
- Important: Even though everyone knows how messages sent to me were encrypted, only I can decrypt them. !!??

- i) Alice wants to send M to Bob
- ii) In public directory, Alice looks up Bob's Public Key, P_B
- iii) Alice sends $P_B(M)$ to Bob
- iv) Bob uses his Secret Key, S_B to decrypt $M = S_B(P_B(M))$

The Black Pages
Public Key Directory





Public-Key Cryptosystems

- No agreement in advance on a secret code.
- Sender (Alice) and receiver (Bob) each have both a public key and a secret key.
- KP_A : Alice's public key; KS_A : Alice's secret key. KP_B : Bob's public key; KS_B : Bob's secret key.
- Public key available to anyone (including eavesdroppers).
 Secret key is kept by owner.
- Functions associated with KS_A , KP_A , KS_B , KP_B are S_A , P_A , S_B , and P_B . S_A and P_A are inverses; S_B and S_B are inverses; So, for any message M

$$M = S_A(P_A(M)) = P_A(S_A(M)),$$

 $M = S_B(P_B(M)) = P_B(S_B(M)).$

An unsecure public key cryptosystem

- Messages are numbers in range 1 to 999.
- Bob's public key is function $P_B(M) = rev(1000 M)$; rev() reverses the digits of a number.
- To encrypt M = 167, Alice sends Bob C = rev(1000 167) = rev(833) = 338.
- In this case, $S_B(C) = 1000 rev(C)$, so Bob can easily decode the message.
- Problem: this is *Not* secure, because *anyone* who knows public key, P_B , can figure out secret key S_B .

Challenge: In order for a public-key cryptosystem to work we must be able to find public/secret key pairs such that

- Receiver Bob can easily calculate $S_B(X)$
- No one else knowing **public key**, P_B , will easily be able to figure out our **secret key**, S_B .

Constructing such **pubic/secret key pairs** sounds almost impossible. Surprisingly, in the mid 1970s, Rivest, Shamir and Adelman, figured out how to do this using simple modular arithmetic.

The result is the **RSA Public Key Cryptosystem**, which is the basis for most e-commerce. We will learn its details in the lecture following the next one.