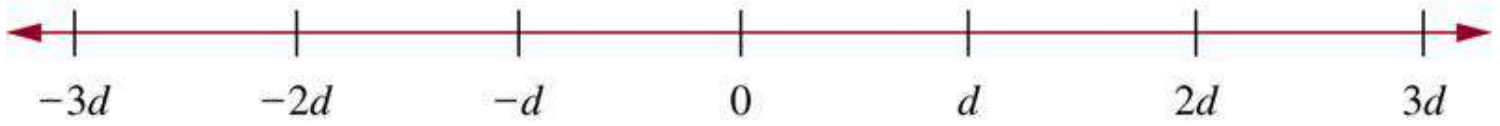
4.1 Divisibility and modular arithmetic

- **Number theory**: the branch of mathematics involves integers and their properties
- If a and b are integers with $a \neq 0$, we say that a divides b if there is an integer c s.t. $b=ac$
- When a divides b we say that a is a **factor** of b and that b is a **multiple** of a
- The notation $a \mid b$ denotes a divides b. We write $a \nmid b$ when does not divide b

Example

- Let n and d be positive integers. How many positive integers not exceeding n are divisible by d?
- The positive integers divisible by d are all integers of them form dk, where k is a positive integer
- Thus, there are $\lfloor n/d \rfloor$ positive integers not exceeding n that are divisible by d



Theorem and corollary

- Theorem: Let a, b, and c be integers, then
  - If a | b and a| c, then a | (b+c)
  - If a | b, and a | bc for all integers c
  - If a | b and b | c, then a | c
- Corollary: If a, b, and c are integers s.t. a | b and a | c, then a | mb+nc whenever m and n are integers

The division algorithm

- Let a be integer and d be a positive integer. Then there are unique integers q and r with $0 \leq r < d$,

  s.t. $a=dq+r$
- In the equality, q is the quotient, r is the remainder

  q = a **div** d, r = a **mod** d
- -11 divided by 3
- $-11=3(-4)+1$, $-4=-11$ div 3, $1=-11$ mod 3
- $-11=3(-3)-2$, but remainder cannot be negative

Modular arithmetic

- If a and b are integers and m is a positive integer, then a is **congruent** to b modulo m if m divides a-b
- We use the notation $a \equiv b \pmod{m}$ to indicate that a is **congruent** to b modulo m
- If a and b are not congruent modulo m, we write $a \not\equiv b \pmod{m}$
- Let a and b be integers, m be a positive integer. Then $a \equiv b \pmod{m}$ if and only if a mod m = b mod m

Example

- Determine whether 17 is congruent to 5 modulo 6, and whether 24 and 14 are not congruent modulo 6
  - 17-5=12, we see $17 \equiv 5 \pmod 6$
  - 24-14=10, and thus $24 \not\equiv 14 \pmod 6$

Theorem

- Karl Friedrich Gauss developed the concept of congruences at the end of 18^{th} century
- Let m be a positive integer. The integer a and b are congruent modulo m if and only if there is an integer k such that a=b+km
  - (→) If a=b+km, then km=a-b, and thus m divides a-b and so a≡b (mod m)
  - (←) if a≡b (mod m), then m | a-b. Thus, a-b=km, and so a=b+km

Theorem

- Let m be a positive integer. If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a+c=b+d \pmod{m}$ and $ac \equiv bd \pmod{m}$
  - Since $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, there are integers s.t. $b=a+sm$ and $d=c+tm$
  - Hence, $b+d=(a+c)+m(s+t)$, $bd=(a+sm)(c+tm)=ac+m(at+cs+stm)$
  - Hence $a+c \equiv b+d \pmod{m}$, and $ac \equiv bd \pmod{m}$

Example

- $7 \equiv 2 \pmod 5$ and $11 \equiv 1 \pmod 5$, so
  - $18 = 7 + 11 \equiv 2 + 1 = 3 \pmod 5$
  - $77 = 7 \cdot 11 \equiv 2 \cdot 1 = 2 \pmod 5$

Corollary

- Let m be a positive integer and let a and b be integers, then

  $(a+b) \bmod m = ((a \bmod m) + (b \bmod m)) \bmod m$

  $ab \bmod m = ((a \bmod m)(b \bmod m)) \bmod m$
- Proof: By definitions mod m and congruence modulo m, we know that $a \equiv (a \bmod m) \pmod m$ and $b \equiv (b \bmod m) \pmod m$. Hence
  - $(a+b) \equiv ((a \bmod m)+(b \bmod m)) \pmod m$
  - $ab \equiv (a \bmod m)(b \bmod m) \pmod m$

4.2 Integer representations and algorithms

- Base b expansion of n
- For instance, $(245)_8=2*8^2+4*8+5=165$
- Hexadecimal expansion of (2AE0B)16

  $(2AE0B)_{16}=2*16^4+10*16^3+14*16^2+0*16+11=175627$
- Constructing base b expansion

Base conversion

- Constructing the base b expansion

  $n=bq_0+a_0$, $0 \leq a_0<b$
- The remainder a_0, is the rightmost digit in the base b expansion of n
- Next, divide q_0 by b to obtain

  $q_0=bq_1+a_1$, $0 \leq a_1<b$
- We see a_1 is the second digit from the right in the base b expansion of n
- Continue this process, successively dividing the quotients by b, until the quotient is zero

Example

- Find the octal base of $(12345)_{10}$
- First, 12345=8*1543+1
- Successively dividing quotients by 8 gives

  1543=8*192+7

  192=8*24+0

  24=8*3+0

  3=8*0+3
- $(12345)_{10}=(30071)_8$

Modular exponentiation

- Need to find b^n mod m efficiently in cryptography
- Impractical to compute b^n and then mod m
- Instead, find binary expansion of n first, e.g., $n=(a_{k-1} \ldots a_1 a_0)$

$$b^n = b^{a_{k-1}\cdot 2^{k-1}+\cdots+a_1\cdot 2+a_0} = b^{a_{k-1}\cdot 2^{k-1}} b^{a_{k-2}\cdot 2^{k-2}} b^{a_1\cdot 2} b^{a_0}$$

- To compute b^n , first find the values of b, b^2, …, $(b^4)^2=b^8$, …
- Next multiple the b^{2^j} where $a_j=1$

Example

- To compute 3^{11}
- $11=(1011)_2$,So $3^{11}=3^8\ 3^2\ 3^1$. First compute $3^2=9$, and then $3^4=9^2=81$, and $3^8=(3^4)^2=(81)^2=6561$, So $3^{11}=6561*9*3=177147$
- The algorithm successively finds b mod m, b^2 mod m, b^4 mod m, …, $b^{2^{k-1}}$ mod m, and multiply together those terms

Algorithm

- **procedure** *modular exponentiation* (*b:integer, n=*($a_{k-1}a_{k-2}a_1a_0$, …, a_n)$_2$, m:positive integer)

  x := 1

  power:=b mod m

  for i:=0 to k-1

  if a_i =1 **then** x:=(x· power) mod m

  power:=(power·power) mod m

  end

  {x equals b^n mod m}

- It uses O((log m)^2 long n) bit operations

Example

- Compute 3^{644} mod 645
  - First note that $644=(1010000100)_2$
  - At the beginning, x=1, power=3 mod 645 = 3
  - i=0, a_0=0, x=1, power=3^2 mod 645=9
  - i=1, a_1=0, x=1, power=9^2 mod 645=81
  - i=2, a_2=1, x=1*81 mod 645=81, power=81^2 mod 645=6561 mod 645=111
  - i=3, a_3=0, x=81, power=111^2 mod 645=12321 mod 645=66
  - i=4, a_4=0, x=81, power=66^2 mod 645=4356 mod 645=486
  - i=5, a_5=0, x=81, power=486^2 mod 645=236196 mod 645=126
  - i=6, a_6=0, x=81, power=126^2 mod 645=15876 mod 645=396
  - i=7, a_7=1, x=(81*396) mod 645=471, power=396^2 mod 645=156816 mod 645=81
  - i=8, a_8=0, x=471, power=81^2 mod 645=6561mod 645=111
  - i=9, a_9=1, x=(471*111) mod 645=36
- 3^{644} mod 645=36

4.3 Primes and greatest common divisions

- **Prime**: a positive integer p greater than 1 if the only positive factors of p are 1 and p
- A positive integer greater than 1 that is not prime is called **composite**
- **Fundamental theorem of arithmetic**: Every positive integer greater than 1 can be written uniquely as a prime or as the product of two or more primes when the prime factors are written in order of non-decreasing size

Example

- Prime factorizations of integers
  - $100=2\cdot2\cdot5\cdot5=2^2\cdot5^2$
  - $641=641$
  - $999=3\cdot3\cdot3\cdot37=3^3\cdot37$
  - $1024=2\cdot2\cdot2\cdot2\cdot2\cdot2\cdot2\cdot2\cdot2\cdot2=2^{10}$

Theorem

- Theorem: If n is a composite integer, then n has a prime division less than or equal to $\sqrt{n}$
- As n is composite, n has a factor 1<a<n, and thus n=ab
- We show that a≤$\sqrt{n}$ or b ≤$\sqrt{n}$ (by contraposition)
- Thus n has a divisor not exceeding $\sqrt{n}$
- This divisor is either prime or by the fundamental theorem of arithmetic, has a prime divisor less than itself, and thus a prime divisor less than less than $\sqrt{n}$
- In either case, n has a prime divisor b ≤ $\sqrt{n}$

Example

- Show that 101 is prime
- The only primes not exceeding $\sqrt{101}$ are 2, 3, 5, 7
- As 101 is not divisible by 2, 3, 5, 7, it follows that 101 is prime

Procedure for prime factorization

- Begin by diving n by successive primes, starting with 2
- If n has a prime factor, we would find a prime factor not exceeding $\sqrt{n}$
- If no prime factor is found, then n is prime
- Otherwise, if a prime factor p is found, continue by factoring n/p
- Note that n/p has no prime factors less than p
- If n/p has no prime factor greater than or equal to p and not exceeding its square root, then it is prime
- Otherwise, if it has a prime factor q, continue by factoring n/(pq)
- Continue until factorization has been reduced to a prime

Example

- Find the prime factorization of 7007
- Start with 2, 3, 5, and then 7, 7007/7=1001
- Then, divide 1001 by successive primes, beginning with 7, and find 1001/7=143
- Continue by dividing 143 by successive primes, starting with 7, and find 143/11=13
- As 13 is prime, the procedure stops
- $7007=7\cdot 7\cdot 11\cdot 13=7^2\cdot 11\cdot 13$

4.3 Theorem

- Theorem: There are infinitely many primes
- Proof by contradiction
- Assume that there are only finitely many primes, p_1, p_2, …, p_n. Let $Q=p_1p_2…p_n+1$
- By Fundamental Theorem of Arithmetic: Q is prime or else it can be written as the product of two or more primes

Theorem

- However, none of the primes p_j divides Q, for if $p_j \mid Q$, then p_j divides $Q-p_1 p_2 \ldots p_n = 1$
- Hence, there is a prime not in the list $p_1, p_2, \ldots, p_n$
- This prime is either Q, if it is prime, or a prime factor for Q
- This is a contradiction as we assumed that we have listed all the primes

Mersenne primes

- As there are infinite number of primes, there is an ongoing quest to find larger and larger prime numbers
- The largest prime known has been an integer of special form 2^p-1 where p is also prime
- Furthermore, currently it is not possible to test numbers not of this or certain other special forms anywhere near as quickly as determine whether they are prime

Mersenne primes

- $2^2-1=3$, $2^3-1=7$, $2^5-1=31$ are Mersenne primes while $2^{11}-1=2047$ is not a Mersenne prime ($2047=23 \cdot 89$)
- Mersenne claims that 2^p-1 is prime for p=2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257 but is composite for all other primes less than 257
  - It took over 300 years to determine it is wrong 5 times
  - For p=67, p=257, 2^p-1 is not prime
  - But p=61, p=87, and p=107, 2^p-1 is prime
- The largest Mersenne prime known (as of early 2011) is $2^{43,112,609}-1$, a number with over 13 million digits

Distribution of primes

- **The prime number theorem**: The ratio of the number of primes not exceeding x and x/ln x approaches 1 as x grows without bound
- Can use this theorem to estimate the odds that a randomly chosen number is prime
- The odds that a randomly selected positive integer less than n is prime are approximately

  $(n/\ln n)/n = 1/\ln n$
- The odds that an integer near 10^{1000} is prime are approximately $1/\ln 10^{1000}$, approximately 1/2300

Open problems about primes

- **Goldbach's conjecture**: every even integer n, n>2, is the sum of two primes

  4=2+2, 6=3+3, 8=5+3, 10=7+3, 12=7+5, …
- As of 2011, the conjecture has been checked for all positive even integers up to $1.6 \cdot 10^{18}$
- **Twin prime conjecture**: Twin primes are primes that differ by 2. There are infinitely many twin primes

Greatest common divisors

- Let a and b be integers, not both zero. The largest integer d such that d | a and d | b is called the **greatest common divisor** (GCD) of a and b, often denoted as gcd(a,b)
- The integers a and b are **relative prime** if their GCD is 1

  gcd(10, 17)=1, gcd(10, 21)=1, gcd(10,24)=2
- The integers a_1, a_2, …, a_n are **pairwise relatively prime** if $gcd(a_i, a_j)=1$ whenever $1 \leq i < j \leq n$

Prime factorization and GCD

- Finding GCD

$$a = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}, b = p_1^{b_1} p_2^{b_2} \cdots p_n^{b_n}$$

$$\gcd(a,b) = p_1^{\min(a_1,b_1)} p_2^{\min(a_2,b_2)} \cdots p_n^{\min(a_n,b_n)}$$

$$120 = 2^3 \cdot 3 \cdot 5, \quad 500 = 2^2 \cdot 5^3$$

$$\gcd(120,500) = 2^2 \cdot 3^0 \cdot 5^1 = 20$$

- **Least common multiples** of the positive integers a and b is the smallest positive integer that is divisible by both a and b, denoted as lcm(a,b)

Least common multiple

- Finding LCM

$$a = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}, b = p_1^{b_1} p_2^{b_2} \cdots p_n^{b_n}$$

$$\operatorname{lcm}(a,b) = p_1^{\max(a_1,b_1)} p_2^{\max(a_2,b_2)} \cdots p_n^{\max(a_n,b_n)}$$

$$120 = 2^3 \cdot 3 \cdot 5, 500 = 2^2 \cdot 5^3$$

$$\operatorname{lcm}(120,500) = 2^3 \cdot 3^1 \cdot 5^3 = 8 \cdot 3 \cdot 125 = 3000$$

- Let a and b be positive integers, then

$ab = \gcd(a,b) \cdot \operatorname{lcm}(a,b)$

Euclidean algorithm

- Need more efficient prime factorization algorithm
- Example: Find gcd(91,287)
- $287 = 91 \cdot 3 + 14$
- Any divisor of 287 and 91 must be a divisor of $287 - 91 \cdot 3 = 14$
- Any divisor of 91 and 14 must also be a divisor of $287 = 91 \cdot 3$
- Hence, the gcd(91,287)=gcd(91,14)
- Next, $91 = 14 \cdot 6 + 7$
- Any divisor of 91 and 14 also divides $91 - 14 \cdot 6 = 7$ and any divisor of 14 and 7 divides 91, i.e., gcd(91,14)=gcd(14,7)
- $14 = 7 \cdot 2$, gcd(14,7)=7, and thus gcd(287,91)=gcd(91,14)=gcd(14,7)=7

Euclidean algorithm

- Lemma: Let a=bq+r, where a, b, q, and r are integers. Then gcd(a,b)=gcd(b,r)
- Proof: Suppose d divides both a and b. Recall if d|a and d|b, then d|a-bk for some integer k. It follows that d also divides a-bq=r. Hence, any common division of a and b is also a common division of b and r
- Suppose that d divides both b and r, then d also divides bq+r=a. Hence, any common divisor of b and r is also common divisor of a and b
- Consequently, gcd(a, b)=gcd(b,r)

Euclidean algorithm

- Suppose a and b are positive integers, a≥b. Let r_0=a and r_1=b, we successively apply the division algorithm

$$
\begin{aligned}
r_0 &= r_1 q_1 + r_2, 0 \le r_2 < r_1 \\
r_1 &= r_2 q_2 + r_3, 0 \le r_3 < r_2 \\
&\cdots \\
r_{n-2} &= r_{n-1} q_{n-1} + r_n, 0 \le r_n < r_{n-1} \\
r_{n-1} &= r_n q_n \\
\gcd(a,b) &= \gcd(r_0, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{n-2}, r_{n-1}) \\
&= \gcd(r_{n-1}, r_n) = \gcd(r_n, 0) = r_n
\end{aligned}
$$

- Hence, the gcd is the last nonzero remainder in the sequence of divisions

Example

- Find the GCD of 414 and 662

  $662 = 414 \cdot 1 + 248$

  $414 = 248 \cdot 1 + 166$

  $248 = 166 \cdot 1 + 82$

  $166 = 82 \cdot 2 + 2$

  $82 = 2 \cdot 41$

  gcd(414,662)=2 (the last nonzero remainder)

$a = bq + r$
$\gcd(a,b) = \gcd(b,r)$

The Euclidean algorithm

- **procedure** *gcd*(*a, b:* positive integers)

  x := a

  y:=b

  while (y≠*0*)

  begin

  r:=x mod y

  x:=y

  y:=r

  end {gcd(a,b)=x}

- The time complexity is O(log b) (where a ≥ b)

4.5 Applications of congruence

- Hashing function: h(k) where k is a key
- One common function: h(k)=k mod m where m is the number of available memory location
- For example, m=111,
  - h(064212848)=064212848 mod 111=14
  - h(037149212)=037149212 mod 111=65
- Not one-to-one mapping, and thus needs to deal with collision
  - h(107405723)=107405723 mod 111 = 14
  - Assign to the next available memory location

Pseudorandom numbers

- Generate random numbers
- The most commonly used procedure is the linear congruential method
  - Modulus m, multiple a, increment c, and seed x_0, with $2 \leq a < m$, $0 \leq c < m$, and $0 \leq x_0 < m$
  - Generate a sequence of pseudorandom numbers $\{x_n\}$ with $0 \leq x_n < m$ for all n, by

    $x_{n+1} = (ax_n + c) \bmod m$

Example

- Let m=9, a=7, c=4, x_0=3
  - $x_1=7x_0+4 \bmod 9=(21+4) \bmod 9=25 \bmod 9 = 7$
  - $x_2=7x_1+4 \bmod 9=(49+4) \bmod 9=53 \bmod 9 = 8$
  - $x_3=7x_2+4 \bmod 9=(56+4) \bmod 9=60 \bmod 9 = 6$
  - $x_4=7x_3+4 \bmod 9=(42+4) \bmod 9=46 \bmod 9 = 1$
  - $x_5=7x_4+4 \bmod 9=(7+4) \bmod 9=11 \bmod 9 = 2$
  - $x_6=7x_5+4 \bmod 9=(14+4) \bmod 9=18 \bmod 9 = 0$
  - $x_7=7x_6+4 \bmod 9=(0+4) \bmod 9=4 \bmod 9 = 4$
  - $x_8=7x_7+4 \bmod 9=(28+4) \bmod 9=32 \bmod 9 =5$
  - $x_9=7x_8+4 \bmod 9=(35+4) \bmod 9=11 \bmod 9 = 3$

$$x_{n+1}=(ax_n+c) \bmod m$$

- A sequence of 3, 7, 8, 6, 1, 2, 0, 4, 5, 3, 7, 8, 6, 1, 2, 0, 4, 5, 3, …
- Contains 9 different numbers before repeating

4.6 Cryptology

- One of the earliest known use is by Julius Caesar, shift each letter by 3

  $f(p)=(p+3) \bmod 26$

  - Translate “meet you in the park”
  - 12 4 4 19 24 14 20 8 13 19 7 4 15 0 17 10
  - 15 7 7 22 1 17 23 11 16 22 10 7 18 3 20 13
  - “phhw brx lq wkh sdun”
  - To decrypt, $f^{-1}(p)=(p-3) \bmod 26$

Example

- Other options: shift each letter by k
  - $f(p)=(p+k) \bmod 26$, with $f^{-1}(p)=(p-k) \bmod 26$
  - $f(p)=(ap+k) \bmod 26$

RSA cryptosystem

- Each individual has an encryption key consisting of a modulus n=pq, where p and q are large primes, say with 200 digits each, and an exponent e that is relatively prime to (p-1)(q-1) (i.e., gcd(e, (p-1)(q-1))=1)
- To transform M: Encryption: $C=M^e \bmod n$, Decryption: $C^d=M \pmod{pq}$
- The product of these primes n=pq, with approximately 400 digits, cannot be factored in a reasonable length of time (the most efficient factorization methods known as of 2005 require billions of years to factor 400-digit integers)