CSE2023 Discrete Computational Structures

Lecture 13

4.3 Theorem

- Theorem: There are infinitely many primes
- Proof by contradiction
- Assume that there are only finitely many primes, $p_1, p_2, \ldots, p_n$. Let $Q=p_1p_2 \ldots p_n+1$
- 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 · 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

5

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

6

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

7

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 ≤ i < j ≤ n

8

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)·lcm(a,b)

Euclidean algorithm

- Need more efficient prime factorization algorithm
- Example: Find gcd(91,287)
- 287=91 · 3 +14
- Any divisor of 287 and 91 must be a divisor of 287- 91 · 3 =14
- Any divisor of 91 and 14 must also be a divisor of 287= 91 · 3
- Hence, the gcd(91,287)=gcd(91,14)
- Next, 91= 14 · 6+7
- Any divisor of 91 and 14 also divides 91- 14 · 6=7 and any divisor of 14 and 7 divides 91, i.e., gcd(91,14)=gcd(14,7)
- 14= 7 · 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 · 1+248

  414=248 · 1+166

  248=166 · 1+82

  166=82 · 2 + 2

  82=2 · 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 $\mathbf{x_0}$, with $2 \le a < m$, $0 \le c < m$, and $0 \le x_0 < m$
  - Generate a sequence of pseudorandom numbers $\{x_n\}$ with $0 \le x_n < m$ for all n, by $x_{n+1}=(ax_n+c) \bmod m$

Example

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

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

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: $C=M^e$ mod n, Decryption: $C^d=M$ (mod 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)

21