Number Theory

Slides by Christopher M. Bourke
Instructor: Berthe Y. Choueiry

Spring 2006

Computer Science & Engineering 235
Introduction to Discrete Mathematics
Sections 2.4–2.6 of Rosen

Introduction I

When talking about division over the integers, we mean division with no remainder.

Definition

Let $a, b \in \mathbb{Z}, a \neq 0$, we say that a *divides* b if there exists $c \in \mathbb{Z}$ such that $b = ac$. We denote this, $a \mid b$ and $a \nmid b$ when a does not divide b. When $a \mid b$, we say a is a *factor* of b.

Theorem

Let $a, b, c \in \mathbb{Z}$ then

1. *If $a \mid b$ and $a \mid c$ then $a \mid (b + c)$.*
2. *If $a \mid b$, then $a \mid bc$ for all $c \in \mathbb{Z}$.*
3. *If $a \mid b$ and $b \mid c$, then $a \mid c$.*

Introduction II

Corollary

If $a, b, c \in \mathbb{Z}$ such that $a \mid b$ and $a \mid c$ then $a \mid mb + nc$ for $n, m \in \mathbb{Z}$.

Division Algorithm I

Let a be an integer and d be a positive integer. Then there are unique integers q and r, with:

- $0 \leq r \leq d$
- such that $a = dq + r$

Not really an algorithm (traditional name). Further:

- a is called the divident
- d is called the divisor
- q is called the quotient
- r is called the remainder, and is positive.

Primes I

Definition

A positive integer $p > 1$ is called *prime* if its only positive factors are 1 and p.
If a positive integer is not prime, it is called *composite*.

Primes II

Theorem (Fundamental Theorem of Arithmetic, FTA)

Every positive integer $n > 1$ can be written uniquely as a prime or as the product of the powers of two or more primes written in nondecreasing size.

That is, for every $n \in \mathbb{Z}, n > 1$, can be written as

$$n = {p_1}^{k_1} {p_2}^{k_2} \cdots {p_l}^{k_l}$$

where each p_i is a prime and each $k_i \geq 1$ is a positive integer.

Given a positive integer, $n > 1$, how can we determine if n is prime or not?

For hundreds of years, people have developed various tests and algorithms for *primality testing*. We'll look at the oldest (and most inefficient) of these.

Lemma

If n is a composite integer, then n has a prime divisor $x \leq \sqrt{n}$.

Proof.

$\square$

Proof.

- Let n be a composite integer.

$\square$

Proof.

- Let n be a composite integer.
- By definition, n has a prime divisor a with $1 < a < n$, thus $n = ab$.

$\square$

Proof.

- Let n be a composite integer.
- By definition, n has a prime divisor a with $1 < a < n$, thus $n = ab$.
- Its easy to see that either $a \leq \sqrt{n}$ or $b \leq \sqrt{n}$. Otherwise, if on the contrary, $a > \sqrt{n}$ *and* $b > \sqrt{n}$, then

$$ab > \sqrt{n}\sqrt{n} = n$$

□

Proof.

- Let n be a composite integer.
- By definition, n has a prime divisor a with $1 < a < n$, thus $n = ab$.
- Its easy to see that either $a \leq \sqrt{n}$ or $b \leq \sqrt{n}$. Otherwise, if on the contrary, $a > \sqrt{n}$ *and* $b > \sqrt{n}$, then
$$ab > \sqrt{n}\sqrt{n} = n$$
- Finally, either a or b is prime divisor or has a factor that is a prime divisor by the Fundamental Theorem of Arithmetic, thus n has a prime divisor $x \leq \sqrt{n}$.

$\square$

Sieve of Eratosthenes
Algorithm

This result gives us an obvious algorithm. To determine if a number n is prime, we simple must test every prime number p with $2 \leq p \leq \sqrt{n}$.

SIEVE

```
INPUT  : A positive integer n ≥ 4.
OUTPUT : true if n is prime.
FOREACH prime number p, 2 ≤ p ≤ √n DO
    IF p | n THEN
        output false
    END
END
output true
```

Can be improved by reducing the upper bound to $\sqrt{\frac{n}{p}}$ at each iteration.

Sieve of Eratosthenes

Efficiency?

This procedure, called the Sieve of Eratosthenes, is quite old, but works.

In addition, it is *very* inefficient. At first glance, this may seem counter intuitive.

This procedure, called the Sieve of Eratosthenes, is quite old, but works.

In addition, it is *very* inefficient. At first glance, this may seem counter intuitive.

- The outer for-loop runs for every prime $p \leq \sqrt{n}$.

Sieve of Eratosthenes

Efficiency?

This procedure, called the Sieve of Eratosthenes, is quite old, but works.

In addition, it is *very* inefficient. At first glance, this may seem counter intuitive.

- The outer for-loop runs for every prime $p \leq \sqrt{n}$.
- Assume that we get such a list *for free*. The loop still executes about

$$\frac{\sqrt{n}}{\ln \sqrt{n}}$$

times (see distribution of primes: next topic, also Theorem 5, page 157).

Sieve of Eratosthenes

Efficiency?

This procedure, called the Sieve of Eratosthenes, is quite old, but works.

In addition, it is *very* inefficient. At first glance, this may seem counter intuitive.

- The outer for-loop runs for every prime $p \leq \sqrt{n}$.
- Assume that we get such a list *for free*. The loop still executes about
$$\frac{\sqrt{n}}{\ln \sqrt{n}}$$
times (see distribution of primes: next topic, also Theorem 5, page 157).
- Assume also that division is our elementary operation.

This procedure, called the Sieve of Eratosthenes, is quite old, but works.

In addition, it is *very* inefficient. At first glance, this may seem counter intuitive.

- The outer for-loop runs for every prime $p \leq \sqrt{n}$.
- Assume that we get such a list *for free*. The loop still executes about
$$\frac{\sqrt{n}}{\ln \sqrt{n}}$$
times (see distribution of primes: next topic, also Theorem 5, page 157).
- Assume also that division is our elementary operation.
- Then the algorithm is $\mathcal{O}(\sqrt{n})$.

This procedure, called the Sieve of Eratosthenes, is quite old, but works.

In addition, it is *very* inefficient. At first glance, this may seem counter intuitive.

- The outer for-loop runs for every prime $p \leq \sqrt{n}$.
- Assume that we get such a list *for free*. The loop still executes about
$$\frac{\sqrt{n}}{\ln \sqrt{n}}$$
times (see distribution of primes: next topic, also Theorem 5, page 157).
- Assume also that division is our elementary operation.
- Then the algorithm is $\mathcal{O}(\sqrt{n})$.
- However, what is the actual *input size*?

- Recall that it is $\log(n)$. Thus, the algorithm runs in *exponential* time with respect to the input size.

- Recall that it is $\log(n)$. Thus, the algorithm runs in *exponential* time with respect to the input size.
- To see this, let $k = \log(n)$

Sieve of Eratosthenes
Efficiency?

- Recall that it is $\log(n)$. Thus, the algorithm runs in *exponential* time with respect to the input size.
- To see this, let $k = \log(n)$
- Then $2^k = n$ and so

$$\sqrt{n} = \sqrt{2^k} = 2^{k/2}$$

- Recall that it is $\log(n)$. Thus, the algorithm runs in *exponential* time with respect to the input size.
- To see this, let $k = \log(n)$
- Then $2^k = n$ and so

$$\sqrt{n} = \sqrt{2^k} = 2^{k/2}$$

- Thus the Sieve is exponential in the input size k.

Sieve of Eratosthenes

Efficiency?

- Recall that it is $\log(n)$. Thus, the algorithm runs in *exponential* time with respect to the input size.
- To see this, let $k = \log(n)$
- Then $2^k = n$ and so

$$\sqrt{n} = \sqrt{2^k} = 2^{k/2}$$

- Thus the Sieve is exponential in the input size k.

The Sieve also gives an algorithm for determining the *prime factorization* of an integer. To date, no one has been able to produce an algorithm that runs in sub-exponential time. The hardness of this problem is the basis of *public-key cryptography*.

Numerous algorithms for primality testing have been developed over the last 50 years.

In 2002, three Indian computer scientists developed the first *deterministic polynomial-time* algorithm for primality testing, running in time $\mathcal{O}(\log^{12}(n))$.

M. Agrawal and N. Kayal and N. Saxena. PRIMES is in P. *Annals of Mathematics*, 160(2):781-793, 2004.

Available at http://projecteuclid.org/Dienst/UI/1.0/Summarize/euclid.annm/1111770735

How Many Primes?

How many primes are there?

Theorem

There are infinitely many prime numbers.

The proof is a simple proof by contradiction.

Proof.

$\square$

Proof.

- Assume to the contrary that there are a finite number of primes, $p_1, p_2, \ldots, p_n$.

$\square$

Proof.

- Assume to the contrary that there are a finite number of primes, $p_1, p_2, \ldots, p_n$.
- Let

$$Q = p_1 p_2 \cdots p_n + 1$$

□

Proof.

- Assume to the contrary that there are a finite number of primes, $p_1, p_2, \ldots, p_n$.
- Let
$$Q = p_1 p_2 \cdots p_n + 1$$
- By the FTA, Q is either prime (in which case we are done) or Q can be written as the product of two or more primes.

$\square$

Proof.

- Assume to the contrary that there are a finite number of primes, $p_1, p_2, \ldots, p_n$.
- Let
$$Q = p_1 p_2 \cdots p_n + 1$$
- By the FTA, Q is either prime (in which case we are done) or Q can be written as the product of two or more primes.
- Thus, one of the primes p_j $(1 \leq j \leq n)$ must divide Q, but then if $p_j \mid Q$, it must be the case that
$$p_j \mid Q - p_1 p_2 \cdots p_n = 1$$

□

Proof.

- Assume to the contrary that there are a finite number of primes, $p_1, p_2, \ldots, p_n$.
- Let
$$Q = p_1 p_2 \cdots p_n + 1$$
- By the FTA, Q is either prime (in which case we are done) or Q can be written as the product of two or more primes.
- Thus, one of the primes p_j $(1 \leq j \leq n)$ must divide Q, but then if $p_j \mid Q$, it must be the case that
$$p_j \mid Q - p_1 p_2 \cdots p_n = 1$$
- Since this is not possible, we've reached a contradiction—there are not finitely many primes.

$\square$

Distribution of Prime Numbers

Theorem

The ratio of the number of prime numbers not exceeding n and $\frac{n}{\ln n}$ approaches 1 as $n \to \infty$.

In other words, for a fixed natural number, n, the number of primes not greater than n is about

$$\frac{n}{\ln n}$$

Mersenne Primes I

A *Mersenne* prime is a prime number of the form

$$2^k - 1$$

where k is a positive integer. They are related to *perfect numbers* (if M_n is a Mersenne prime, $\frac{M_n(M_n+1)}{2}$ is perfect).

Perfect numbers are numbers that are equal to the sum of their proper factors, for example $6 = 1 \cdot 2 \cdot 3 = 1 + 2 + 3$ is perfect.

Mersenne Primes II

It is an open question as to whether or not there exist odd perfect numbers. It is also an open question whether or not there exist an infinite number of Mersenne primes.

Such primes are useful in testing suites (i.e., benchmarks) for large super computers.

To date, 42 Mersenne primes have been found. The last was found on February 18th, 2005 and contains 7,816,230 digits.

Division

Theorem (The Division "Algorithm")

Let $a \in \mathbb{Z}$ and $d \in \mathbb{Z}^+$ then there exists unique integers q, r with $0 \leq r < d$ such that

$$a = dq + r$$

Some terminology:

- d is called the *divisor*.
- a is called the *dividend*.
- q is called the *quotient*.
- r is called the *remainder*.

We use the following notation:

$$\begin{aligned} q &= a \textbf{ div } d \\ r &= a \textbf{ mod } d \end{aligned}$$

Greatest Common Divisor I

Definition

Let a and b be integers not both zero. The largest integer d such that $d \mid a$ and $d \mid b$ is called the *greatest common divisor* of a and b. It is denoted

$$\gcd(a, b)$$

The $\gcd$ is always guaranteed to exist since the set of common divisors is finite. Recall that 1 is a divisor of any integer. Also, $\gcd(a, a) = a$, thus

$$1 \leq \gcd(a, b) \leq \min\{a, b\}$$

Greatest Common Divisor II

Definition

Two integers a, b are called *relatively prime* if

$$\gcd(a, b) = 1$$

Sometimes, such integers are called *coprime*.

There is natural generalization to a set of integers.

Definition

Integers $a_1, a_2, \ldots, a_n$ are *pairwise relatively prime* if $\gcd(a_i, a_j) = 1$ for $i \neq j$.

The $\gcd$ can "easily"[1] be found by finding the prime factorization of two numbers.

Let

$$\begin{aligned} a &= {p_1}^{a_1}{p_2}^{a_2}\cdots{p_n}^{a_n} \\ b &= {p_1}^{b_1}{p_2}^{b_2}\cdots{p_n}^{b_n} \end{aligned}$$

Where each power is a nonnegative integer (if a prime is not a divisor, then the power is 0).

Then the $\gcd$ is simply

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

[1] Easy *conceptually*, not computationally

Example

What is the $\gcd(6600, 12740)$?
The prime decompositions are

$$\begin{aligned} 6600 &= 2^3 3^1 5^2 7^0 11^1 13^0 \\ 12740 &= 2^2 3^0 5^1 7^2 11^0 13^1 \end{aligned}$$

So we have

$$\begin{aligned} \gcd(6600, 12740) &= 2^{\min\{2,3\}} 3^{\min\{0,1\}} 5^{\min\{1,2\}} 7^{\min\{0,2\}} \\ &\quad 11^{\min\{0,1\}} 13^{\min\{0,1\}} \\ &= 2^2 3^0 5^1 7^0 11^0 13^0 \\ &= 20 \end{aligned}$$

Least Common Multiple

Definition

The *least common multiple* of positive integers a, b is the smallest positive integer that is divisible by both a and b. It is denoted

$$\mathrm{lcm}(a, b)$$

Again, the lcm has an "easy" method to compute. We still use the prime decomposition, but use the $\max$ rather than the $\min$ of powers.

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

Example

What is the $\mathrm{lcm}(6600, 12740)$?
Again, the prime decompositions are

$$\begin{aligned} 6600 &= 2^3 3^1 5^2 7^0 11^1 13^0 \\ 12740 &= 2^2 3^0 5^1 7^2 11^0 13^1 \end{aligned}$$

So we have

$$\begin{aligned} \mathrm{lcm}(6600, 12740) &= 2^{\max\{2,3\}} 3^{\max\{0,1\}} 5^{\max\{1,2\}} 7^{\max\{0,2\}} \\ &\quad 11^{\max\{0,1\}} 13^{\max\{0,1\}} \\ &= 2^3 3^1 5^2 7^2 11^1 13^1 \\ &= 4,204,200 \end{aligned}$$

Intimate Connection

There is a very close connection between the $\gcd$ and lcm.

Theorem

Let $a, b \in \mathbb{Z}^+$, *then*

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

Proof?

Often, rather than the quotient, we are only interested in the remainder of a division operation. We introduced the notation before, but we formally define it here.

Definition

Let $a, b \in \mathbb{Z}$ and $m \in \mathbb{Z}^+$. Then a is *congruent to* b *modulo* m if m divides $a - b$. We use the notation

$$a \equiv b (\bmod\ m)$$

If the congruence does not hold, we write $a \not\equiv b (\bmod\ m)$

An equivalent characterization can be given as follows.

Theorem

Let $m \in \mathbb{Z}^+$. *Then* $a \equiv b \pmod{m}$ *if and only if there exists* $q \in \mathbb{Z}$ *such that*

$$a = qm + b$$

i.e. a quotient q.

Alert: $a, b \in \mathbb{Z}$, i.e. can be negative or positive.

Theorem

Let $a, b \in \mathbb{Z}, m \in \mathbb{Z}^+$. *Then,*

$$a \equiv b \pmod{m} \iff a \textbf{ mod } m = b \textbf{ mod } m$$

Theorem

Let $m \in \mathbb{Z}^+$. *If* $a \equiv b \pmod{m}$ *and* $c \equiv d \pmod{m}$ *then*

$$a + c \equiv b + d \pmod{m}$$

and

$$ac \equiv bd \pmod{m}$$

Example

- $36 \equiv 1 \pmod{5}$ since the remainder of $\frac{36}{5}$ is 1.

Example

- $36 \equiv 1 \pmod 5$ since the remainder of $\frac{36}{5}$ is 1.
- Similarly, $-17 \equiv -1 \pmod 2$, $-17 \equiv 1 \pmod 2$, $-17 \equiv 3 \pmod 2$, etc.

Example

- $36 \equiv 1 \pmod{5}$ since the remainder of $\frac{36}{5}$ is 1.
- Similarly, $-17 \equiv -1 \pmod{2}$, $-17 \equiv 1 \pmod{2}$, $-17 \equiv 3 \pmod{2}$, etc.
- However, we prefer to express congruences with $0 \leq b < m$.

Example

- $36 \equiv 1 \pmod 5$ since the remainder of $\frac{36}{5}$ is 1.
- Similarly, $-17 \equiv -1 \pmod 2$, $-17 \equiv 1 \pmod 2$, $-17 \equiv 3 \pmod 2$, etc.
- However, we prefer to express congruences with $0 \leq b < m$.
- $64 \equiv 0 \pmod 2$, $64 \equiv 1 \pmod 3$, $64 \equiv 4 \pmod 5$, $64 \equiv 4 \pmod 6$, $64 \equiv 1 \pmod 7$, etc.

Inverses I

Definition

An *inverse* of an element x modulo m is an integer x^{-1} such that

$$xx^{-1} \equiv 1 (\bmod\ m)$$

Inverses do not always exist, take $x = 5, m = 10$ for example.

The following is a necessary and sufficient condition for an inverse to exist.

Theorem

Let a and m be integers, $m > 1$. A (unique) inverse of a modulo m exists if and only if a and m are relatively prime.