Lecture 7:
Number Theory

Steven Skiena

Department of Computer Science
State University of New York
Stony Brook, NY 11794–4400

http://www.cs.sunysb.edu/∼skiena

Number Theory and Divisibility

> "G-d created the integers. All else is the work of man."
> – Kronecker.

Number theory is the study of the properties of the integers, specifically integer divisibility. It is a fascinating area is beautiful proofs and surprising results.

We say b *divides* a (denoted $b|a$) if $a = bk$ for some integer k. Equivalently, we say that b is a *divisor of* a or a is a *multiple* of b if $b|a$.

As a consequence of this definition, the smallest natural divisor of every non-zero integer is 1. In general there is no integer k such that $a = 0 \cdot k$.

Prime Numbers

A *prime number* is an integer $p > 1$ which is only divisible by 1 and itself.

Said another way, if p is a prime number, then $p = a \cdot b$ for integers $a \leq b$ implies that $a = 1$ and $b = p$.

The *fundamental theorem of arithmetic*. states is that every integer can be expressed in only one way as the product of primes. 105 is uniquely expressed as $3 \times 5 \times 7$, and 32 is uniquely expressed as $2 \times 2 \times 2 \times 2 \times 2$.

This unique set of numbers multiplying to n is called the *prime factorization* of n.

Any number which is not prime is said to be *composite*.

Primality Testing and Factorization

Not only are there an infinite number of primes (see Euclid's proof), but there are a lot of them. There are roughly $x/\ln x$ primes less than or equal to x, i.e. roughly one out of every $\ln x$ numbers is prime.

The easiest way to find the prime factorization of an integer n is repeated division. Note the that smallest prime factor is must at most $\sqrt{n}$ unless n is prime.

Constructing all Divisors

Every divisor is the product of some subset of these prime factors. Such subsets can be constructed using backtracking techniques, but we must be careful about duplicate prime factors.

For example, the prime factorization of 12 has three terms (2, 2, and 3) but 12 has only 6 divisors (1, 2, 3, 4, 6, 12).

Greatest Common Divisor

The *greatest common divisor*, or *gcd*, the *largest* divisor shared by a given pair of integers. The reduced form of the fraction $24/36$ comes after we divide both the numerator and denominator by $gcd(x, y)$, in this case 12.
Euclid's GCD algorithm rests on two observations. First,

> If $b|a$, then $gcd(a, b) = b$.

This should be pretty clear. If b divides a, then $a = bk$ for some integer k, and thus $gcd(bk, b) = b$. Second,

> If $a = bt + r$ for integers t and r, then $gcd(a, b) = gcd(b, r)$.

Why? By definition, $gcd(a, b) = gcd(bt + r, b)$. Any common divisor of a and b must rest totally with r, because bt clearly must be divisible by any divisor of b.
It can also find integers x and y such that

$$a \cdot x + b \cdot y = gcd(a, b)$$

which will prove quite useful in solving linear congruences.

Implementation

```
/*      Find the gcd(p,q) and x,y such that p*x + q*y = gcd(p,q)      */

long gcd(long p, long q, long *x, long *y)
{
        long x1,y1;                     /* previous coefficients */
        long g;                         /* value of gcd(p,q) */

        if (q > p) return(gcd(q,p,y,x));

        if (q == 0) {
                *x = 1;
                *y = 0;
                return(p);
        }

        g = gcd(q, p%q, &x1, &y1);

        *x = y1;
        *y = (x1 - floor(p/q)*y1);

        return(g);
}
```

The *least common multiple* (lcm), the *smallest* integer which is divided by both of a given pair of integers. For example, the least common multiple of 24 and 36 is 72.
To compute it, observe that $lcm(x, y) = xy/gcd(x, y)$.

Modular Arithmetic

Sometimes computing the remainder is more important than a quotient.

What day of the week will your birthday fall on next year? All you need to know is the remainder of the number of days between now and then (either 365 or 366) when dividing by the 7 days of the week. Thus it will fall on this year's day plus one ($365 \bmod 7$) or two ($366 \bmod 7$) days, depending upon whether it is affected by a leap year.

The key to such efficient computations is *modular arithmetic*. The number we are dividing by is called the *modulus*, and the remainder left over is called the *residue*.

What is $(x+y)\text{mod}n$? We can simplify this to

$$((x\text{mod}n)+(y\text{mod}n))\text{mod}n$$

to avoid adding big numbers.
Subtraction s just a special case of addition.

$$(12\text{mod}100)-(53\text{mod}100)=-41\text{mod}100=59\text{mod}100$$

Notice how we can convert a negative number mod n to a positive number by adding a multiple of n to it.

Exponentiation

Since multiplication is just repeated addition,

$$xy \text{mod } n = (x \text{mod } n)(y \text{mod } n) \text{mod } n$$

Since exponentiation is just repeated multiplication,

$$x^y \text{mod} n = (x \text{mod} n)^y \text{mod} n$$

Since exponentiation is the quickest way to produce really large integers, this is where modular arithmetic really proves its worth.

Division proves considerably more complicated to deal with...

What is the Last Digit?

What is the last digit of 2^{100}? We can do this computation by hand. What we really want to know is what $2^{100}\text{mod}10$ is. By doing repeated squaring, and taking the remainder mod10 at each step we make progress very quickly:

$$
\begin{aligned}
2^{3}\text{mod}10 &= 8 \\
2^{6}\text{mod}10 &= 8 \times 8\text{mod}10 \rightarrow 4 \\
2^{12}\text{mod}10 &= 4 \times 4\text{mod}10 \rightarrow 6 \\
2^{24}\text{mod}10 &= 6 \times 6\text{mod}10 \rightarrow 6 \\
2^{48}\text{mod}10 &= 6 \times 6\text{mod}10 \rightarrow 6 \\
2^{96}\text{mod}10 &= 6 \times 6\text{mod}10 \rightarrow 6 \\
2^{100}\text{mod}10 &= 2^{96} \times 2^{3} \times 2^{1}\text{mod}10 \rightarrow 6
\end{aligned}
$$

Congruences

Congruences are an alternate notation for representing modular arithmetic. We say that $a \equiv b (\text{mod} m)$ if $m|(a-b)$. By definition, if $a \text{mod} m$ is b, then $a \equiv b (\text{mod } m)$.

It gets us thinking about the *set* of integers with a given remainder n, and gives us equations for representing them.

What integers x satisfy the congruence $x \equiv 3 (\text{mod } 9)$? The set of solutions is all integers of the form $9y + 3$, where y is any integer.

What about $2x \equiv 3 (\text{mod } 9)$ and $2x \equiv 3 (\text{mod } 4)$?

Trial and error should convince you that exactly the integers of the form $9y + 6$ satisfy the first example, while the second has no solutions at all.

Operations on Congruences

Congruences support addition, subtraction, and multiplication, as well as a limited form of division – provided they share the same modulus:

- *Addition and Subtraction* – Suppose $a \equiv b (\text{mod } n)$ and $c \equiv d (\text{mod } n)$. Then $a + c \equiv b + d (\text{mod } n)$. If $4x \equiv 7 (\text{mod } 9)$ and $3x \equiv 3 (\text{mod } 9)$, then

$$4x - 3x \equiv 7 - 3 (\text{mod } 9) \rightarrow x \equiv 4 (\text{mod } 9)$$

- *Multiplication* – General multiplication holds, i.e., $a \equiv b (\text{mod } n)$ and $c \equiv d (\text{mod } n)$ implies $ac \equiv bd (\text{mod } n)$.

- *Division* – We cannot cavalierly cancel common factors from congruences. Note that $6 \cdot 2 \equiv 6 \cdot 1 (\text{mod } 3)$, but clearly $2 \not\equiv 1 (\text{mod } 3)$.

Note that division can be defined as multiplication by an inverse, so a/b is equivalent to ab^{-1}. But this inverse does not always exist – try to find a solution to $2x \equiv 1 (\text{mod } 4)$.

Simplifying and Solving Congruences

We *can* simplify a congruence $ad \equiv bd (\text{mod } dn)$ to $a \equiv b (\text{mod } n)$, so we can divide all three terms by a mutually common factor if one exists. Thus $170 \equiv 30 (\text{mod } 140)$ implies that $17 \equiv 3 (\text{mod } 14)$.

A linear congruence is an equation of the form $ax \equiv b (\text{mod } n)$. Solving this equation means identifying which values of x satisfy it.

Not all such equations have solutions. $ax \equiv 1 (\text{mod } n)$ has a solution if and only if the modulus and the multiplier are relatively prime, i.e., $gcd(a, n) = 1$. We may use Euclid's algorithm to find this inverse through the solution to $a \cdot x' + n \cdot y' = gcd(a, n) = 1$.

In general, there are three cases, depending on the relationship between a, b, and n:

- $\gcd(a, b, n) > 1$ – Then we can divide all three terms by this divisor to get an equivalent congruence. This gives us a single solution mod the new base, or equivalently $\gcd(a, b, n)$ solutions $(\text{mod } n)$.
- $gcd(a, n)$ *does not divide* b – The congruence can have no solution.
- $gcd(a, n) = 1$ – Then there is one solution $(\text{mod } n)$. Further, $x = a^{-1}b$ works, since $aa^{-1}b \equiv b(\text{mod } n)$. This inverse exists and can be found using Euclid's algorithm.

More Advanced Tools

The *Chinese remainder theorem* gives us a tool for working with systems of congruences over different moduli. Suppose there is exists an integer x such that $x \equiv a_1 (\text{mod } m_1)$ and $x \equiv a_2 (\text{mod } m_2)$. Then x is uniquely determined $(\text{mod } m_1 m_2)$ if m_1 and m_2 are relatively prime.

Diophantine Equations

Diophantine equations are formulae in which the variables are restricted to integers. For example, Fermat's last theorem concerned answers to the equation $a^n + b^n = c^n$.

The most important class are linear Diophantine equations of the form $ax - ny = b$, where x and y are the integer variables and a, b, and n are integer constants. These are readily shown to be equivalent to the solving the congruence $ax \equiv b (\mathrm{mod}\, n)$ and hence can be solved using the techniques above.

110702 (Carmichael Numbers)

Which composite integers n always satisfy the equation

$$a^n \text{mod} n = a$$

Does this require large integer arithmetic?

110704 (Factovisors)

Does a divide $n!$?

Does this require large integer arithmetic?

110707 (Marbles)

What is the cheapest way to perfectly store n among two types of boxes?

110708 (Repackaging)

Cups of three sizes are sold in certain predefined sets with a given number of each type of cup per package.
Can we buy packages so we end up with exactly the same number of all three types of cups?