GB13604 - Maths for Computer Science

Lecture 3 – Number Theory

Claus Aranha

caranha@cs.tsukuba.ac.jp

College of Information Science

2018-10-17

Last updated October 14, 2018

Summary Week 1 and 2

- Proof by Cases
- Proof by Contradiction (Well Ordered Principle)
- Proof by Induction
- Sets Definition
- Sets Relationships
- Finite Set Sizes

Exercise Discussion

For This Lecture...

Number Theory – Textbook Chapter 8

- GCD and Extended GCD
- Modular Arithmetic, and Relatively Primes
- Euler's Theorem, and Rings
- RSA Algorithm

Some basic arithmetic assumptions

For the proofs in this class, we can assume some default rules for arithmetic operators: *, +, -, ...

- $a(b + c) = ab + ac$
- $ab = ba$
- $a(bc) = (ab)c$
- $a + 0 = a$
- $a - a = 0$
- $a + 1 > a$
- etc...

The Division Theorem

Axiom:
For any $b > 0$ and a in $\mathbb{N}$, we have:

- q = quotient(a,b)
- r = remainder(a,b)

$\exists$ **unique** q and r in $\mathbb{N}$ such as

$$a = bq + r, 0 < r \leq a$$

Take this by granted too!

Divisibility

c **divides** $a (c|a)$ iff

$$\exists k, a = k \times c.$$

- $5|15$ because $15 = 3 \times 5$
- $n|0$ because $0 = 0 \times n$
- $1|n$ because $n = n \times 1$

Simple Divisibility Facts

- $c|a$ implies $c|(sa)$
  $a = kc$ implies $(sa) = (sk)c$ multiply s on both sides
- $c|a$ and $c|b$ implies $c|(a+b)$
  $a = k_1 c, b = k_2 c, a + b = k_1 c + k_2 c = (k_1 + k_2)c$
- $c|a$ and $c|b$ implies $c|(sa+tb)$
  $sa + tb$ is a linear combination of a and b

  This one is pretty important!

Common Divisors

If $c|a$ and $c|b$, then c is a common divisor of a and b.

Common divisors of a and b also divide linear combinations of a and b.

Greatest Common Divisor

We define $gcd(a, b) ::=$ the greatest **common divisor** of a and b.

- $gcd(10, 12) = 2$ $(10 = 2 \times 5, 12 = 2 \times 6)$
- $gcd(13, 12) = 1$ No common factors and $1|x, \forall x$
- $gcd(17, 17) = 17$
- $gcd(0, n) = n$ for $n > 0$

Does **one** gcd aways exists? (Yes, because of the Well Ordering Principle)

Greatest Common Divisor

We define $gcd(a, b) ::=$ the greatest **common divisor** of a and b.

- **lemma**: p is prime implies that gcd(p,a) = 1 or p;
- **proof**: The only divisors of p are ± 1 and $\pm p$.

Euclidean Algorithm (GCD is easy to compute)

Remainder Lemma: gcd(a,b) = gcd(b, rem(a,b)) for $b \neq 0$

Proof idea:

- $a = qb + r, 0 \leq r < b$ (division axiom)
- Any divisor of two out of $\{a, qb, r\}$, must divide all three. (Check this yourself using slide 8)
- Therefore, $\forall m$ if $m|a$ and $m|b$ then $m|$rem(a,b)

Example GCD (Using Remainder Lemma)

GCD(899, 493) – a = 899, b = 493

- $899 = 493 \times 1 + 406$ — division axiom
- GCD(899, 493) = GCD(493, 406) — remainder lemma
- GCD(493, 406) = GCD(406, 87) — $493 = 406 \times 1 + 87$
- GCD(406,87) = GCD(87,58) — $406 = 87 \times 4 + 58$
- GCD(87,58) = GCD(58,29) = GCD(29,0) = 29

This is a **fast** algorith (proof later)

GCD as a State Machine

- States::= $\mathbb{N} \times \mathbb{N}$
- Start State::= (a, b)
- State Transitions::= $(x, y) \rightarrow (y, rem(x, y))$ for $y \neq 0$

GCD as a State Machine

Proof of Partial Correctness

1. We want to show: P((x,y))::= [gcd(x,y) = gcd(a,b)]
2. P(start) is trivially true: (gcd(a,b) = gcd(a,b))
3. P is a Preserved Invariant:
   GCD(x,y) = GCD(y,rem(x,y)) (remainder lemma)
4. By 2 and 3, P holds for any state in the machine.
5. So if the machine stops, $x = gcd(a, b)$. Why?
   - The machine only stops when $y = 0$
   - GCD(x,0) = x

GCD as a State Machine

Proof of Termination

- At each transition, y is replaced with rem(x,y)
- $0 < \text{rem(x,y)} \leq y$. (division axiom)
- So eventually $y = 0$, and the machine halts.

- how fast does it halt?
- At each transition, x is replaced by y. Two cases:
  - $y \leq x/2$ so x is halved this step.
  - $y > x/2$ so rem(x,y) = x - y < (x/2), so x gets halved at the next step.
- x gets halved (or even smaller) every two steps.
- So number of steps is $\leq 2\log_2 b$

GCD and Linear Combinations

Extended Euclid Algorithm or The Pulverizer

Main Idea:

- GCD(a,b) is a linear combination of a and b.
- GCD(a,b) $= sa + tb$.
- **collorary:** All lin. comb. of a,b are multiples of GCD(a,b)

- The Pulverizer helps us find s and t

The Pulverizer: Method

Calculate euclid's algorithm:

- GCD(x,y) = GCD(y,rem(x,y) **Start**: GCD(a,b)

Keep track of four coefficient: **c,d,e,f**

- $x = ca + db$ and $y = ea + fb$
- **at start:** x = 1a + 0b, y = 0a+1b
- **update:** $x_{\text{next}} = y = ea + fb$
- $y_{\text{next}} = \text{rem}(x, y) = x - qy = ca + db - q(ea + fb)$
- $y_{\text{next}} = (c - qe)a + (d - qf)b$

The Pulverizer: Example

a = 899, b = 493

hfill (remember: $e_1 = c_0 - q_0 e_0, f_1 = d_0 - q_0 f_0$)

a	b	q	rem(a,b)	c	d	e	f
899	493	1	406	1	0	0	1
493	406	1	87	0	1	1	-1
406	87	4	58	1	-1	-1	2
87	58	1	29	-1	2	5	-9
58	29	2	0	5	-9	-6	11
29	0	-	-	-6	11	-	-

GCD(899,493) = 29 = $-6 \times 899 + 11 \times 493$

The Pulverizer: One Weird Trick

$$\mathrm{GCD}(899, 493) = -6 \times 899 + 11 \times 493$$

How can I get a positive coefficient for 899?

$$\mathrm{GCD}(899, 493) = (-6+493k)899+(11-899k)493, \text{ for any } k$$

Let $k = 1$

$$\mathrm{GCD}(899, 493) = 487 \times 899 - 888 \times 493$$

Remember Robot 1.0?

- It could move 5 steps forward, 3 steps back.
- How many moves it takes to reach “8”?
- GCD(5,3) $= 1 = 2 \times 5 - 3 \times 3$
- $8 = 8 \times 1 = (8 \times 2)5 - (8 \times 3)3$
- 16 steps forward, 24 steps back.

- Not the most efficient solution, but we can find any solution with this strategy.

Prime Factorization Theorem

- Lemma: if p prime and p|ab, then p|a or p|b
- **Proof**: suppose **not(p|a)**, then GCD(p,a) = 1
- So: $\exists s, t. sa + tp = 1$, multiply everything by b
- sab + tbp = b
- $p|sab$ and $p|tbp$, so $p|(sab + tbp)$ and $p|b$ **done.**

- Corolary: if $p|a_1 a_2 \ldots a_m$ then $\exists i. p|a_i$
- **proof**: Induction on m

Prime Factorization Theorem

Fundamental Theorem of Arithmetic

Every Integer > 1 factors uniquely into a weakly decreasing sequence of primes.

$$n > 1, \qquad n = p_1 p_2 p_3 \ldots p_k, \qquad p_1 \geq p_2 \geq \ldots \geq p_k$$

Example

$61394323221 = 53 \times 37 \times 37 \times 37 \times 11 \times 11 \times 7 \times 3 \times 3 \times 3$

Prime Factorization Theorem

Proof by Contradiction.

- Suppose $n > 1$ does not have a unique prime factorization (it can be factored in two different ways).
- By WOP, there is a minimal n where theorem is false.
- $n = p_1 p_2 p_3 \dots p_k$ and $n = q_1 q_2 q_3 \dots q_{k'}$
- if $p_1 = q_1$ then we can cancel them, and n is not smallest anymore. ($n' = p_2 \dots p_k = q_2 \dots q_{k'}$)
- **So we assume** $q_1 > p_1$
- **By the corolary** $q_1 | n \to q_1 | p_i \in p_1 p_2 \dots p_k$
- But, because $q_1 > p_i \forall i$, this is impossible. **done.**

Congruences mod N

Congruence mod n: Definition

$a \equiv b \pmod{n}$ iff $n|(a-b)$

Examples:

- $30 \equiv 12 \pmod{9}$ because 9|(30-12)
- $66666663 \equiv 788253 \pmod{10}$

Congruence has many applications in crypto and hashing.

Remainder Theorem

$a \equiv b$ (mod n) **iff** rem(a,n) = rem(b,n)

(This is the CS "a%n" definition)

Proof: (rem(a,b) = $r_{a,b}$)

- Let $a = q_a n + r_{a,n}$, $b = q_b n + r_{b,n}$
- **if** $r_{a,n} = r_{b,n}$ then $a - b = (q_a - q_b)n \rightarrow n|(a-b)$
- **also if** $n|(a-b)$ then $n|((q_a - q_b)n + (r_{a,n} - r_{b,n}))$
- but $0 \leq r_{*,n} < n$ so $r_{a,n} - r_{b,n}$ must be 0

Remainder Theorem: Consequences

$a \equiv b$ (mod n) means that rem(a,n) = rem(b,n).

Consequences:

- $a \equiv b$ (mod n) **implies that** $b \equiv a$ (mod n)
- $a \equiv b$ (mod n) **and** $b \equiv c$ (mod n) **implies** $a \equiv c$ (mod n)
- $a \equiv$rem(a,n) (mod n) (important!)
- **If** $a \equiv b$ (mod n) **then** $a + c \equiv b + c$ (mod n)
- **If** $a \equiv b$ (mod n) **then** $ac \equiv bc$ (mod n)
- **If** $a \equiv b$ (mod n) **and** $c \equiv d$ (mod n)
  then $a + c \equiv b + d$ (mod n) **and** $ac \equiv bd$ (mod n)

What does this mean?

Overall, arithmetic (mod n) is very similar to normal arithmetic.

If $a \equiv a'$ (mod n) and a' is simpler, you can usually replace in the formula to make it easier.

Using $a \equiv$ rem(a,n) (mod n) means that we can keep the numbers in modular arithmetic between 0 and n.

Modular Arithmetic: Example

- What is $287^9 \equiv ? \pmod 4$
- $287^9 \equiv 3^9 \pmod 4$ **because** $r_{287,4} = 3$
- $3^9 = ((3^2)^2)^2 \times 3$
- $((3^2)^2)^2 \times 3 \equiv (1^2)^2 \times 3 \pmod 4$ **because** $9 \equiv 1 \pmod 4$
- $289^9 \equiv 3 \pmod 4$

- And we did not need to calculate any x^9!

Difference between Arithmetic and Modular Arithmetic

We saw that Arithmetic and Modular Arithmetic are similar but...

- $8 \times 2 \equiv 3 \times 2 \pmod{10}$
- **Can we do:** $8 \times \cancel{2} \equiv 3 \times \cancel{2} \pmod{10}$?
- $8 \not\equiv 3 \pmod{10}$

- We can't cancel arbitrarily!

When can we cancel $ak \equiv bk \pmod{n}$?

You can cancel when k and n have no common factors.

OR, when GCD(k,n) = 1

Modular Inverses

- Modular Inverse: If GCD(k,n) = 1 then $\exists k', k \times k' \equiv 1 \pmod{n}$
- If $ak \equiv bk \pmod{n}$, we can multiply both sides by k'
- $akk' \equiv bkk' \pmod{n} \rightarrow 1a \equiv 1b \pmod{n}$

k has an inverse (mod n) **iff** k is relatively prime to n

Euler's Function

Number of relatively primes of n between 0 and n

$$\Phi(n) ::= \#k \in [0, n), GCD(k, n) = 1$$

Let us define:

$$\text{gcd1}\{n\} ::= \{k \in [0, n) | \text{GCD}(k, n) = 1\}$$

- $\text{gcd1}\{7\} = \{1, 2, 3, 4, 5, 6\}$ $\Phi(7) = 6$
- $\text{gcd1}\{12\} = \{1, 5, 7, 11\}$ $\Phi(12) = 4$

Calculating $\Phi(n)$

- If n is prime, $\Phi(n) = n - 1$
- If n is a power of a prime, $\Phi(p^k) = p^k - p^{k-1}$
  - Ex: $\Phi(9) = 3^2 - 3 = 6$ $\quad\{1, 2, 4, 5, 7, 8\}$
- If n is ab where GCD(a,b)=1, $\Phi(ab) = \Phi(a)\Phi(b)$
  - Ex: $\Phi(12) = \Phi(3) \times \Phi(4) = (3 - 1) \times (2^2 - 2) = 4$

- Euler's Theorem: if GCD(k,n) = 1, $k^{\Phi(n)} \equiv 1$ (mod n)

The Ring of $\mathbb{Z}_n$

Working with just Remainders

- The integer interval $[0, n)$ under $+, \times(\mathbb{Z}_n)$ is called $\mathbb{Z}_n$.

- $i + j(\mathbb{Z}_n) ::= \mathrm{rem}(i + j, n)$
- $i \times j(\mathbb{Z}_n) ::= \mathrm{rem}(i \times j, n)$

Arithmetic in $\mathbb{Z}_n$

- $3 + 6 = 2(\mathbb{Z}_7)$
- $9 \times 8 = 6(\mathbb{Z}_{11})$
- $\mathrm{rem}(a, n)$ is equivalent to $r(a)(\mathbb{Z}_n)$

$\equiv$ (mod n) and $\mathbb{Z}_n$

$i \equiv j \pmod{n}$ **iff** $r(i) = r(j)(\mathbb{Z}_n)$

As we saw before, most arithmetic rules apply to $\mathbb{Z}_n$ arithmetic.

No Cancelling Rule – Be careful that you cannot easily cancel multiplication!

$$8 \times \cancel{2} \neq 3 \times \cancel{2} (\mathbb{Z}_{10})$$

$\mathbb{Z}_n^*$ – Elements relatively prime to n

- $i \in \mathbb{Z}_n^*$ iff $\gcd(i, n) = 1$
- i is cancellable in $\mathbb{Z}_n$
- i has an inverse in $\mathbb{Z}_n$

- $\Phi(n) ::= |\mathbb{Z}_n^*|$
- Euler's Theorem: $k^{\Phi(n)} = 1(\mathbb{Z}_n)$ if $k \in \mathbb{Z}_n^*$

The RSA Encryption System

- Public Key Cryptosystem;
- Anyone can send a secret (encrypted) message to the receiver without prior contact, using only public information.
- This sounds paradoxical: How can someone construct a secret message using only public information?

RSA Cryptosystem: Basic Assumption

- Basic Assumption: **One Way Functions** that are easy to compute but hard to invert

- It is easy to compute the product n of two large primes p and q ($n = pq$)
- It is very hard to factor n into p and q.

RSA Cryptosystem: Preparations

- sender wants to send a message to receiver
- rcv generates primes p, q, $n ::= pq$
- rcv finds e rel. prime to $(p-1)(q-1)$
  (hint: $(p-1)(q-1) = \Phi(n)$)
- (e,n) ::= public key. rcv publishes it widely.
- rcv finds $d ::= e^{-1}(\mathbb{Z}^*_{(p-1)(q-1)})$
- d ::= private key, rcv keeps it.

RSA Cryptosystem: Message

- sender encodes a message $m \in [1, n)$
- sender reads (e,n) and calculates $\hat{m} = m^e (\mathbb{Z}_n)$
- sender sends $\hat{m}$ to rcv
- rcv calculates $\hat{m}^d = m(\mathbb{Z}_n)$
- Euler's Theorem guarantees that $\hat{m}^d = m, d = e^{-1}, (\mathbb{Z}_n)$

RSA Cryptosystem: Requirements

- Find two large primes, p and q
  - Ok because: Lots of Primes
  - Need fast primality tester
- Find e relatively prime to $(p-1)(q-1)$
  - Ok because: Lots of relatively prime numbers
  - Fast because $\text{GCD}(e, (p-1)(q-1))$ is fast
- Find $e^{-1}(\mathbb{Z}^*_{(p-1)(q-1)})$
  - Fast because of the Pulverizer
- Check the book for the proofs.

Summary of the Class

- GCD algorithm (with proof) and Pulverizer
- Arithmetic modulo n, and $\mathbb{Z}$ ring
- Euler's Theorem
- The RSA cryptosystem

Extra Reading

- Proof for Euler's Theorem
- Relationship between SAT and factoring