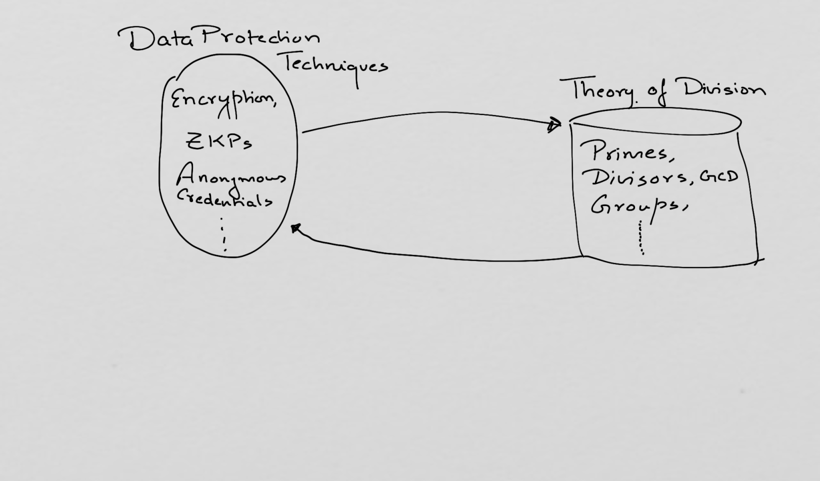
Introduction to Modular Arithmetic

Subodh Sharma

Subhashis Banerjee

IIT Delhi, Computer Science Department

Detour


Data Protection Techniques
Encryption,
ZKPs
Anonymous Credentials
Theory of Division
Primes,
Divisors, GCD
Groups,

Theory of Division

- Let $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ be the set of integers.

Theory of Division

- Let $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ be the set of integers.
- Division Theorem: For any integer a and any positive integer n, there exist unique integers q and r s.t.: $a = qn + r$ where $0 \leq r < n$.

Theory of Division

- Let $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ be the set of integers.
- Division Theorem: For any integer a and any positive integer n, there exist unique integers q and r s.t.: $a = qn + r$ where $0 \leq r < n$.
  - Given $a, n \in \mathbb{Z}, n > 0$ the notation $r = a \ mod \ n$ represents the remainder of the division of a by n and $q = \lfloor a/n \rfloor$ represents the quotient of the division.

Theory of Division

- Let $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ be the set of integers.
- Division Theorem: For any integer a and any positive integer n, there exist unique integers q and r s.t.: $a = qn + r$ where $0 \leq r < n$.
  - Given $a, n \in \mathbb{Z}, n > 0$ the notation $r = a \ mod \ n$ represents the remainder of the division of a by n and $q = \lfloor a/n \rfloor$ represents the quotient of the division.
- $[a]_n$ be the equivalence classes according to the remainders modulo n.

Theory of Division

- Let $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ be the set of integers.
- Division Theorem: For any integer a and any positive integer n, there exist unique integers q and r s.t.: $a = qn + r$ where $0 \leq r < n$.
  - Given $a, n \in \mathbb{Z}, n > 0$ the notation $r = a \; mod \; n$ represents the remainder of the division of a by n and $q = \lfloor a/n \rfloor$ represents the quotient of the division.
- $[a]_n$ be the equivalence classes according to the remainders modulo n.
  - Formally, $[a]_n = \{a + kn : k \in \mathbb{Z}\}$

Theory of Division

- Let $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ be the set of integers.
- Division Theorem: For any integer a and any positive integer n, there exist unique integers q and r s.t.: $a = qn + r$ where $0 \leq r < n$.
    - Given $a, n \in \mathbb{Z}, n > 0$ the notation $r = a \ mod \ n$ represents the remainder of the division of a by n and $q = \lfloor a/n \rfloor$ represents the quotient of the division.
- $[a]_n$ be the equivalence classes according to the remainders modulo n.
    - Formally, $[a]_n = \{a + kn : k \in \mathbb{Z}\}$
    - Example: $[3]_7 = \{\cdots, -11, -4, 3, 10, 17, \cdots\} = [-4]_7 = [10]_7$
    - Thus, $a \in [b]_n$ is the same as writing $a \equiv b \ (mod \ n)$.

Theory of Division

- Let $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ be the set of integers.
- Division Theorem: For any integer a and any positive integer n, there exist unique integers q and r s.t.: $a = qn + r$ where $0 \leq r < n$.
  - Given $a, n \in \mathbb{Z}, n > 0$ the notation $r = a \; mod \; n$ represents the remainder of the division of a by n and $q = \lfloor a/n \rfloor$ represents the quotient of the division.
- $[a]_n$ be the equivalence classes according to the remainders modulo n.
  - Formally, $[a]_n = \{a + kn : k \in \mathbb{Z}\}$
  - Example: $[3]_7 = \{\cdots, -11, -4, 3, 10, 17, \cdots\} = [-4]_7 = [10]_7$
  - Thus, $a \in [b]_n$ is the same as writing $a \equiv b \; (mod \; n)$.
- Set of all such equivalence classes: $\mathbb{Z}_n = \{[a]_n : 0 \leq a \leq n-1\}$ will be read as $\{0, 1, \cdots, n-1\}$

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.
- From above $\Rightarrow d|a+b, d|a-b$. In general, $d|ax+by$ for any integers x and y.

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.
- From above $\Rightarrow d|a+b, d|a-b$. In general, $d|ax+by$ for any integers x and y.
- Greatest Common Divisor (gcd): Among all the *common* divisors of a and b, the largest among them is the $gcd(a, b)$

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.
- From above $\Rightarrow d|a+b, d|a-b$. In general, $d|ax+by$ for any integers x and y.
- Greatest Common Divisor (gcd): Among all the *common* divisors of a and b, the largest among them is the $gcd(a, b)$
  - Eg: $gcd(24, 30) = 6$.
    - $24 = 2.2.2.3$
    - $30 = \quad 2.3.5$

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.
- From above $\Rightarrow d|a+b, d|a-b$. In general, $d|ax+by$ for any integers x and y.
- Greatest Common Divisor (gcd): Among all the *common* divisors of a and b, the largest among them is the $gcd(a,b)$
  - Eg: $gcd(24,30) = 6$.
    - $24 = 2.2.2.3$
    - $30 = \quad 2.3.5$
  - A useful characterization of gcd: If a and b are nonzero, then $gcd(a,b)$ is the smallest positive number of the set $\{ax+by : x, y \in \mathbb{Z}\}$.

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.
- From above $\Rightarrow d|a+b, d|a-b$. In general, $d|ax+by$ for any integers x and y.
- Greatest Common Divisor (gcd): Among all the *common* divisors of a and b, the largest among them is the $gcd(a,b)$
  - Eg: $gcd(24,30) = 6$.
    - $24 = 2.2.2.3$
    - $30 = \quad 2.3.5$
  - A useful characterization of gcd: If a and b are nonzero, then $gcd(a,b)$ is the smallest positive number of the set $\{ax+by : x,y \in \mathbb{Z}\}$.
  - If $gcd(a,b) = 1$, then a and b are relatively prime.

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.
- From above $\Rightarrow d|a+b, d|a-b$. In general, $d|ax+by$ for any integers x and y.
- Greatest Common Divisor (gcd): Among all the *common* divisors of a and b, the largest among them is the $gcd(a,b)$
  - Eg: $gcd(24,30) = 6$.
    - $24 = 2.2.2.3$
    - $30 = \quad 2.3.5$
  - A useful characterization of gcd: If a and b are nonzero, then $gcd(a,b)$ is the smallest positive number of the set $\{ax+by : x,y \in \mathbb{Z}\}$.
  - If $gcd(a,b) = 1$, then a and b are relatively prime.
  - Relatively prime integers: $gcd(a,b) = 1$

Common Divisors

- If $d|a$ and $d|b$, we say d is a common divisor of a and b.
- From above $\Rightarrow d|a+b, d|a-b$. In general, $d|ax+by$ for any integers x and y.
- Greatest Common Divisor (gcd): Among all the *common* divisors of a and b, the largest among them is the $gcd(a,b)$
  - Eg: $gcd(24,30) = 6$.
    - $24 = 2.2.2.3$
    - $30 = \quad 2.3.5$
  - A useful characterization of gcd: If a and b are nonzero, then $gcd(a,b)$ is the smallest positive number of the set $\{ax+by : x,y \in \mathbb{Z}\}$.
  - If $gcd(a,b) = 1$, then a and b are relatively prime.
  - Relatively prime integers: $gcd(a,b) = 1$
- Note: No efficient solution for integer factorization.

Euclid's Greatest Common Divisor Algorithm

- Euclid in his "The Elements" (c. 300 BC) gave a recursive algorithm: $gcd(a, b) = gcd(b, a \ mod \ b)$
  - Let $d = gcd(a, b)$. Then $d \mid a, d \mid b$.
  - $a \ mod \ b = a - qb$ where $q = \lfloor a/b \rfloor$. Thus, $d \mid a \ mod \ b$
  - Similarly, can be shown that $a \ mod \ b \mid d$
- Eg:

$$
\begin{aligned}
gcd(30, 21) &= gcd(21, 9) \\
&= gcd(9, 3) \\
&= gcd(3, 0)
\end{aligned}
$$

Extended Euclid's Algorithm

$d = gcd(a, b) = ax + by$.

- The algorithm solves for x and y. Note that x and y can be zero or negative.
- As efficient as $gcd(a, b)$ computation
- Required to compute modular multiplicative inverses.

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.
  - Eg: $(\mathbb{Z}_n, +_n)$ (Check?)

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.
  - Eg: $(\mathbb{Z}_n, +_n)$ (Check?)
  - What about $(\mathbb{Z}_n, *_n)$ – Answer is No!
    - Not all elements have an inverse! Eg: $0, 2, 3, 4 \in \mathbb{Z}_6$ have no inverses.

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.
  - Eg: $(\mathbb{Z}_n, +_n)$ (Check?)
  - What about $(\mathbb{Z}_n, *_n)$ – Answer is No!
    - Not all elements have an inverse! Eg: $0, 2, 3, 4 \in \mathbb{Z}_6$ have no inverses.
  - Let $\mathbb{Z}_n^* = \{[a]_n : gcd(a, n) = ax + by = 1, x, y \in \mathbb{Z}\}$. Then $(\mathbb{Z}_n^*, *_n)$ is a finite group. Eg: $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$, $\mathbb{Z}_{15}^* = \{1, 2, 4, 7, 8, 11, 13, 14\}$

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.
  - Eg: $(\mathbb{Z}_n, +_n)$ (Check?)
  - What about $(\mathbb{Z}_n, *_n)$ – Answer is No!
    - Not all elements have an inverse! Eg: $0, 2, 3, 4 \in \mathbb{Z}_6$ have no inverses.
  - Let $\mathbb{Z}_n^* = \{[a]_n : gcd(a, n) = ax + by = 1, x, y \in \mathbb{Z}\}$. Then $(\mathbb{Z}_n^*, *_n)$ is a finite group. Eg: $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$, $\mathbb{Z}_{15}^* = \{1, 2, 4, 7, 8, 11, 13, 14\}$
  - Thus, for $a \in \mathbb{Z}_n^*$, we have $ax \equiv 1 \ mod \ n$.
  - This multiplicative inverse x can quickly computed using *Extended Euclid*.

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.
  - Eg: $(\mathbb{Z}_n, +_n)$ (Check?)
  - What about $(\mathbb{Z}_n, *_n)$ – Answer is No!
    - Not all elements have an inverse! Eg: $0, 2, 3, 4 \in \mathbb{Z}_6$ have no inverses.
  - Let $\mathbb{Z}_n^* = \{[a]_n : gcd(a, n) = ax + by = 1, x, y \in \mathbb{Z}\}$. Then $(\mathbb{Z}_n^*, *_n)$ is a finite group. Eg: $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$, $\mathbb{Z}_{15}^* = \{1, 2, 4, 7, 8, 11, 13, 14\}$
  - Thus, for $a \in \mathbb{Z}_n^*$, we have $ax \equiv 1 \ mod \ n$.
  - This multiplicative inverse x can quickly computed using *Extended Euclid*.
  - Eg: $a = 5, n = 11$. Then $(d, x, y) = Extended\_Euclid(a, n) = (1, -2, 1)$. Thus, the multiplicative inverse of 5 is $[-2]_{11}$ or $[9]_1 1$.

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.
  - Eg: $(\mathbb{Z}_n, +_n)$ (Check?)
  - What about $(\mathbb{Z}_n, *_n)$ – Answer is No!
    - Not all elements have an inverse! Eg: $0, 2, 3, 4 \in \mathbb{Z}_6$ have no inverses.
  - Let $\mathbb{Z}_n^* = \{[a]_n : gcd(a, n) = ax + by = 1, x, y \in \mathbb{Z}\}$. Then $(\mathbb{Z}_n^*, *_n)$ is a finite group. Eg: $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$, $\mathbb{Z}_{15}^* = \{1, 2, 4, 7, 8, 11, 13, 14\}$
  - Thus, for $a \in \mathbb{Z}_n^*$, we have $ax \equiv 1 \ mod \ n$.
  - This multiplicative inverse x can quickly computed using *Extended Euclid*.
  - Eg: $a = 5, n = 11$. Then $(d, x, y) = Extended\_Euclid(a, n) = (1, -2, 1)$. Thus, the multiplicative inverse of 5 is $[-2]_{11}$ or $[9]_1 1$.
- Observe: $|\mathbb{Z}_n^*| < |\mathbb{Z}_n|$ when n is composite. Why?

Modular Arithmetic

Finite groups

- Finite Group: $(S, \oplus)$ where S is finite.
  - Properties:
    1. Closure: For all $a, b \in S$, then $a \oplus b \in S$
    2. Associativity: For all $a, b, c \in S$, we have $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
    3. Identity: There exists $e \in S$, s.t. $a \oplus e = e \oplus a = a$, for all $a \in S$.
    4. Inverse: For each $a \in S$, there exists a unique $b \in S$ s.t. $a \oplus b = b \oplus a = e$.
  - Eg: $(\mathbb{Z}_n, +_n)$ (Check?)
  - What about $(\mathbb{Z}_n, *_n)$ – Answer is No!
    - Not all elements have an inverse! Eg: $0, 2, 3, 4 \in \mathbb{Z}_6$ have no inverses.
  - Let $\mathbb{Z}_n^* = \{[a]_n : gcd(a, n) = ax + by = 1, x, y \in \mathbb{Z}\}$. Then $(\mathbb{Z}_n^*, *_n)$ is a finite group. Eg: $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$, $\mathbb{Z}_{15}^* = \{1, 2, 4, 7, 8, 11, 13, 14\}$
  - Thus, for $a \in \mathbb{Z}_n^*$, we have $ax \equiv 1 \ mod \ n$.
  - This multiplicative inverse x can quickly computed using *Extended Euclid*.
  - Eg: $a = 5, n = 11$. Then $(d, x, y) = Extended\_Euclid(a, n) = (1, -2, 1)$. Thus, the multiplicative inverse of 5 is $[-2]_{11}$ or $[9]_11$.
- Observe: $|\mathbb{Z}_n^*| < |\mathbb{Z}_n|$ when n is composite. Why?
- In practice, we choose $\mathbb{Z}_p^*$ where p is prime.

Modular Arithmetic: Subgroups

- Given $(S, \oplus)$, choose any $a \in S$ and compute $a^{(k)} = \underbrace{a \oplus a \oplus \cdots \oplus a}_{\text{k times}}$

Modular Arithmetic: Subgroups

- Given $(S, \oplus)$, choose any $a \in S$ and compute $a^{(k)} = \underbrace{a \oplus a \oplus \cdots \oplus a}_{\text{k times}}$
- The subgroup generated by a is denoted as $(\langle a \rangle, \oplus)$ where $\langle a \rangle = \{a^{(k)} : k \geq 1\}$

Modular Arithmetic: Subgroups

- Given $(S, \oplus)$, choose any $a \in S$ and compute $a^{(k)} = \underbrace{a \oplus a \oplus \cdots \oplus a}_{\text{k times}}$
- The subgroup generated by a is denoted as $(\langle a \rangle, \oplus)$ where $\langle a \rangle = \{a^{(k)} : k \geq 1\}$
- Eg: $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$. Choose $a = 2$. Then $a^{(1)} = 2, a^{(2)} = 4, a^{(3)} = 0, \cdots$ (since $\oplus = +$). For $\mathbb{Z}_6$, we have:
  $\langle 1 \rangle = \{0, 1, 2, 3, 4, 5\}$
  $\langle 2 \rangle = \{0, 2, 4\}$
  $\langle 3 \rangle = \{0, 3\}$

Modular Arithmetic: Subgroups

- Given $(S, \oplus)$, choose any $a \in S$ and compute $a^{(k)} = \underbrace{a \oplus a \oplus \cdots \oplus a}_{\text{k times}}$
- The subgroup generated by a is denoted as $(\langle a \rangle, \oplus)$ where $\langle a \rangle = \{a^{(k)} : k \geq 1\}$
- Eg: $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$. Choose $a = 2$. Then $a^{(1)} = 2, a^{(2)} = 4, a^{(3)} = 0, \cdots$ (since $\oplus = +$). For $\mathbb{Z}_6$, we have:
  $\langle 1 \rangle = \{0, 1, 2, 3, 4, 5\}$
  $\langle 2 \rangle = \{0, 2, 4\}$
  $\langle 3 \rangle = \{0, 3\}$
- Eg: $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$. Choose $a = 2$. Then $a^{(1)} = 2, a^{(2)} = 4, a^{(3)} = 1, \cdots$ (since $\oplus = *$).
  $\langle 1 \rangle = \{1\}$
  $\langle 2 \rangle = \{1, 2, 4\}$
  $\langle 3 \rangle = \{1, 2, 3, 4, 5, 6\}$

Modular Arithmetic (Continued): Subgroups

- *Lagrange's Theorem*: If $(S', \oplus)$ forms a subgroup of $(S, \oplus)$, then $|S'|$ divides $|S|$.
- For a subgroup $\langle a \rangle$, the element a is called the *generator*.

Modular Arithmetic (Continued): Subgroups

- *Lagrange's Theorem*: If $(S', \oplus)$ forms a subgroup of $(S, \oplus)$, then $|S'|$ divides $|S|$.
- For a subgroup $\langle a \rangle$, the element a is called the *generator*.
- Notice that there exists t s.t. $a^{(t)} = e$. The smallest possible t is called the order of $\langle a \rangle$ – denoted by $ord(a)$.

Modular Arithmetic (Continued): Subgroups

- *Lagrange's Theorem*: If $(S', \oplus)$ forms a subgroup of $(S, \oplus)$, then $|S'|$ divides $|S|$.
- For a subgroup $\langle a \rangle$, the element a is called the *generator*.
- Notice that there exists t s.t. $a^{(t)} = e$. The smallest possible t is called the order of $\langle a \rangle$ – denoted by $ord(a)$.
  - for group operaton $+ : a^{(t)} = e$ is the same as $at \equiv 0 \; mod \; n$

Modular Arithmetic (Continued): Subgroups

- *Lagrange's Theorem*: If $(S', \oplus)$ forms a subgroup of $(S, \oplus)$, then $|S'|$ divides $|S|$.
- For a subgroup $\langle a \rangle$, the element a is called the *generator*.
- Notice that there exists t s.t. $a^{(t)} = e$. The smallest possible t is called the order of $\langle a \rangle$ – denoted by $ord(a)$.
  - for group operaton $+ : a^{(t)} = e$ is the same as $at \equiv 0 \; mod \; n$
  - for group operaton $* : a^{(t)} = e$ is the same as $a^t \equiv 1 \; mod \; n$

Modular Arithmetic (Continued): Subgroups

- *Lagrange's Theorem*: If $(S', \oplus)$ forms a subgroup of $(S, \oplus)$, then $|S'|$ divides $|S|$.
- For a subgroup $\langle a \rangle$, the element a is called the *generator*.
- Notice that there exists t s.t. $a^{(t)} = e$. The smallest possible t is called the order of $\langle a \rangle$ – denoted by $ord(a)$.
  - for group operaton $+ : a^{(t)} = e$ is the same as $at \equiv 0 \ mod \ n$
  - for group operaton $* : a^{(t)} = e$ is the same as $a^t \equiv 1 \ mod \ n$
- If $(S, \oplus)$ is a finite group with identity e, then for all $a \in S$: $a^{(|S|)} = e$

Modular Arithmetic (Continued): Subgroups

- *Lagrange's Theorem*: If $(S', \oplus)$ forms a subgroup of $(S, \oplus)$, then $|S'|$ divides $|S|$.
- For a subgroup $\langle a \rangle$, the element a is called the *generator*.
- Notice that there exists t s.t. $a^{(t)} = e$. The smallest possible t is called the order of $\langle a \rangle$ – denoted by $ord(a)$.
  - for group operaton $+ : a^{(t)} = e$ is the same as $at \equiv 0 \; mod \; n$
  - for group operaton $* : a^{(t)} = e$ is the same as $a^t \equiv 1 \; mod \; n$
- If $(S, \oplus)$ is a finite group with identity e, then for all $a \in S$: $a^{(|S|)} = e$
  - Proof from *Lagrange's Theorem* that $ord(a) \mid |S|$

Modular Linear Equations

Consider $ax \equiv b \; mod \; n$, where $a, n > 0$.

- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \; mod \; n : x > 0\}$.

Modular Linear Equations

Consider $ax \equiv b \; mod \; n$, where $a, n > 0$.

- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \; mod \; n : x > 0\}$.
- Thus, the above equation has a solution if and only if $[b] \in \langle a \rangle$.

Modular Linear Equations

Consider $ax \equiv b \ mod \ n$, where $a, n > 0$.

- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \ mod \ n : x > 0\}$.
- Thus, the above equation has a solution if and only if $[b] \in \langle a \rangle$.
  - Precise characterisation: $\langle a \rangle = \langle d \rangle = \{0, d, 2d, \cdots, (n/d - 1)d\}$, where $d = gcd(a, n)$. Thus, $|\langle a \rangle| = n/d$.

Modular Linear Equations

Consider $ax \equiv b \ mod \ n$, where $a, n > 0$.

- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \ mod \ n : x > 0\}$.
- Thus, the above equation has a solution if and only if $[b] \in \langle a \rangle$.
  - Precise characterisation: $\langle a \rangle = \langle d \rangle = \{0, d, 2d, \cdots, (n/d - 1)d\}$, where $d = gcd(a, n)$. Thus, $|\langle a \rangle| = n/d$.
  - $ax \equiv b \ mod \ n$ is solvable for x if and only if $d \mid b, d = gcd(a, n)$.

Modular Linear Equations

Consider $ax \equiv b \ mod \ n$, where $a, n > 0$.

- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \ mod \ n : x > 0\}$.
- Thus, the above equation has a solution if and only if $[b] \in \langle a \rangle$.
  - Precise characterisation: $\langle a \rangle = \langle d \rangle = \{0, d, 2d, \cdots, (n/d - 1)d\}$, where $d = gcd(a, n)$. Thus, $|\langle a \rangle| = n/d$.
  - $ax \equiv b \ mod \ n$ is solvable for x if and only if $d \mid b, d = gcd(a, n)$.
- Either has d distinct solutions where $d = gcd(a, n) \wedge d|b$ or has no solution.

Modular Linear Equations

Consider $ax \equiv b \ mod \ n$, where $a, n > 0$.

- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \ mod \ n : x > 0\}$.
- Thus, the above equation has a solution if and only if $[b] \in \langle a \rangle$.
  - Precise characterisation: $\langle a \rangle = \langle d \rangle = \{0, d, 2d, \cdots, (n/d - 1)d\}$, where $d = gcd(a, n)$. Thus, $|\langle a \rangle| = n/d$.
  - $ax \equiv b \ mod \ n$ is solvable for x if and only if $d \mid b, d = gcd(a, n)$.
- Either has d distinct solutions where $d = gcd(a, n) \wedge d|b$ or has no solution.
- $8x \equiv 2 \ mod \ 12$. Any solution?

Modular Linear Equations

Consider $ax \equiv b \ mod \ n$, where $a, n > 0$.

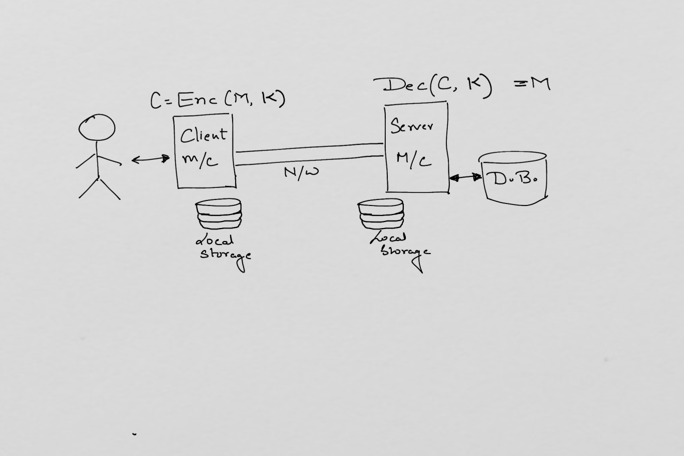
- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \ mod \ n : x > 0\}$.
- Thus, the above equation has a solution if and only if $[b] \in \langle a \rangle$.
  - Precise characterisation: $\langle a \rangle = \langle d \rangle = \{0, d, 2d, \cdots, (n/d - 1)d\}$, where $d = gcd(a, n)$. Thus, $|\langle a \rangle| = n/d$.
  - $ax \equiv b \ mod \ n$ is solvable for x if and only if $d \mid b, d = gcd(a, n)$.
- Either has d distinct solutions where $d = gcd(a, n) \wedge d|b$ or has no solution.
- $8x \equiv 2 \ mod \ 12$. Any solution?
- when $d = 1 \Rightarrow$ the above equation has a *unique* solution.

Modular Linear Equations

Consider $ax \equiv b \ mod \ n$, where $a, n > 0$.

- Choose an $a \in \mathbb{Z}_n$. Then $\langle a \rangle = \{ax \ mod \ n : x > 0\}$.
- Thus, the above equation has a solution if and only if $[b] \in \langle a \rangle$.
  - Precise characterisation: $\langle a \rangle = \langle d \rangle = \{0, d, 2d, \cdots, (n/d - 1)d\}$, where $d = gcd(a, n)$. Thus, $|\langle a \rangle| = n/d$.
  - $ax \equiv b \ mod \ n$ is solvable for x if and only if $d \mid b, d = gcd(a, n)$.
- Either has d distinct solutions where $d = gcd(a, n) \wedge d|b$ or has no solution.
- $8x \equiv 2 \ mod \ 12$. Any solution?
- when $d = 1 \Rightarrow$ the above equation has a *unique* solution.
- Of special interest: $b = 1$ (*multiplicative inverse* of a)

Symmetric Key Encryption



- The same key k is used for Encryption and decryption key
- Encryption produces ciphertext $C = E(M, k)$. Decryption recovers the message $M = D(E(M, k), k)$

Symmetric Key Encryption (Continued)

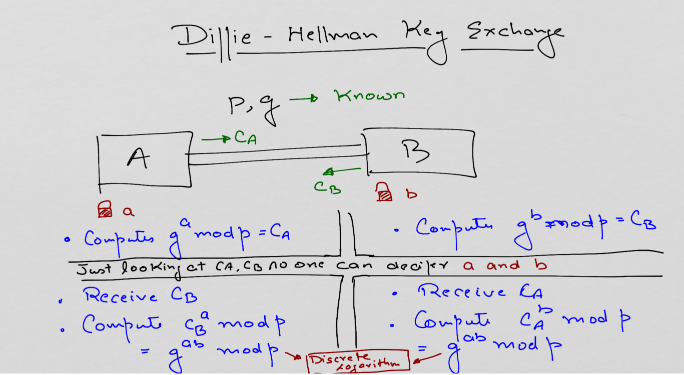
- Substitution ciphers as encryption functions: Cipher alphabet shifted, reversed, or scrambled (Eg: Caesar cipher)
  - MEETME $\rightarrow$ LOOQ LO
  - Security is weak: Frequency distribution of ciphertext which can allow formation of partial words. O is used 3 times. In English, top letters that are frequent used are E, T, A etc. Replacing O with E gives a partial word.
- Similarly, for Transposition cipher: Sliding alphabet of ciphertexts to look for anagrams. Then search the space of anagrams.
- Need to rely on a key whose detection is hard - prime factorisation of large semi-primes is presumably hard!
- Known Symmetric encryption algorithms: AES, 3DES, Blowfish.
- AES128: Runs in 16 rounds. Each round has substitution, permutation, linear transformation, XOR with round key.

More on Symmetric Encryption

- How to securely share the secret key among each pair of communicating parties?
  - Solution: Diffie-Hellman key exdchange protocol.
- After receiving the secret key, how to securely *store* them? Threats from *insider attacks*, compromised privileged software (such as the OS).
  - Note that no data protection technique via key-based data encryption will be adequate without a solution to the secure key storage problem.
- The number of keys to be maintained by each machine is $O(n)$ (where n is the number of machines that it will communicate).

Diffie-Hellman Key Exchange Protocol

- Security of the protocol is derived from the presumed hardness of the *discrete logarithm* problem.
- Protocol begins by choosing a publicly agreed upon a large prime p and the associated primitive root g.
  - Recall that primitive root is that special element $g \in \mathbb{Z}_p^*$ such that $\langle g \rangle = \mathbb{Z}_p^*$.
- Two participants A and B, then choose secret keys a and b, respectively.



Diffie-Hellman Key Exchange Protocol

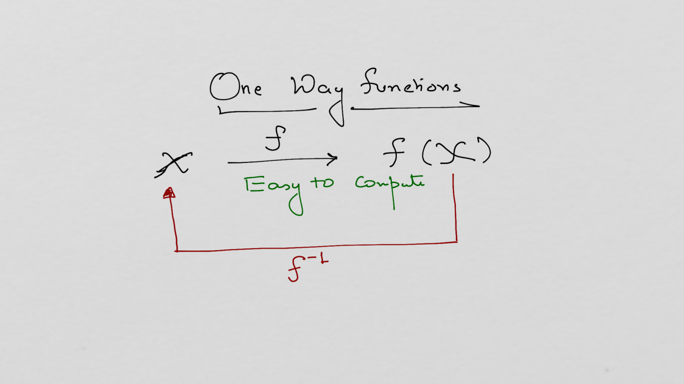
- Participant A computes a ciphertext $C_A = g^a \ (mod\ p)$. Similarly, B computes $C_B = g^b \ (mod\ p)$.
- Participant A sends C_A to participant B and receives C_B from B.
- A computes $C_B^a \ (mod\ p) = g^{ab} \ (mod\ p)$ and B computes $C_A^b \ (mod\ p) = g^{ab} \ (mod\ p)$.
- Thus, the secret key $g^{ab} \ (mod\ p)$ is established.
- Any intruder wishing to read the message will have to find the value ab (*i.e.*, solving the discrete logarithm problem).

Discrete Logarithm Problem

Let us focus on $\mathbb{Z}_n^*$ instead of $\mathbb{Z}_n$.

- We know that for all $a \in \mathbb{Z}_n^*$, $a^{|\mathbb{Z}_n^*|} \equiv 1 \ (mod\ n), n > 1$
  - This is also called *Euler's Theorem*
  - The Euler Phi function is defined as: $\phi(n) = |\mathbb{Z}_n^*|$
- Remember from earlier discussion that $|\mathbb{Z}_p^*| = p - 1$ when p is a prime.
- From Fermat's Theorem: $a^{p-1} \equiv 1 \ (mod\ p)$ for all $a \in \mathbb{Z}_p^*$
- Let $g \in \mathbb{Z}_n^*$ such that $\langle g \rangle = \mathbb{Z}_n^*$. Then $\mathbb{Z}_n^*$ is called *cyclic*.
- By definition of $\langle g \rangle$, for all $a \in \mathbb{Z}_n^*$, there exists z s.t. $g^z \equiv a \ (mod\ n)$.
  - z is called the *discrete logarithm* of a modulo n.

One Way Functions



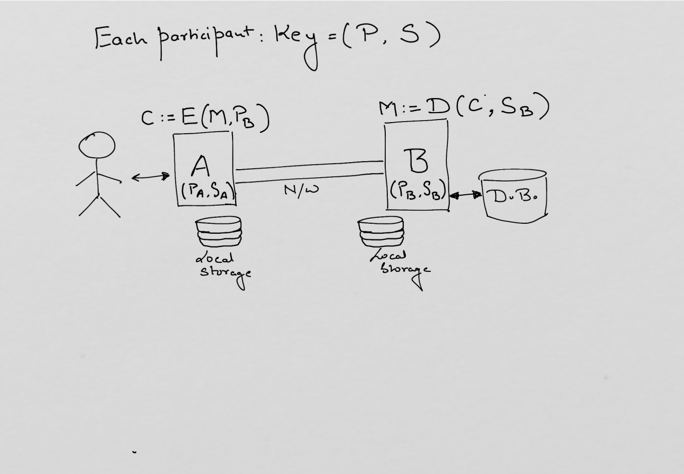
- Given x, computing $F(x)$ is fast.
- However, given $F(x)$, computing $F^{-1}(x)$ is difficult
- Discrete logarithm problem is an instance of a one-way function! That is given g, z, n computing $g^z(mod\ n)$ is fast. But given g, n, a computing $log_g(a)$

One-way Hash Functions

$F(\langle$msg-arbitrary-size$\rangle) = \langle$msg-fixed-size$\rangle$

- Properties:
  - Deterministic: same message produces the same hash.
  - Collision-resistant: It is hard to find two inputs m_1, m_2 s.t. $m_1 \neq m_2$ but $F(m_1) = F(m_2)$.
  - Avalanche effect: A small change in message leads to a large change in the hashed message
- Used in digital signatures, MACs. Egs: SHA-256, MD5
- Security: Brute-force search, Caching the o/p of hash functions (called rainbow table attack).
  - Use of *salt* (a random data as an additional input to the hash function) makes the attack infeasible.

Public-key Cryptosystems



- Every participant computes and maintains a key.
- Each key has two parts: public P, secret S

Public-key Cryptosystems(Cont.)

- Thus, machine A's key is (P_A, S_A) and B's key is (P_B, S_B).
- With a slight abuse of notation we will consider $E(M, P_x)$ in the figure as $P_x(M)$ and $D(M, S_x)$ as $S_x(M)$.
- Public and secret keys are "matched pairs", in the sense that they specify functions that are inverses of each other, *i.e.*, $S_x(P_x(msg)) = P_x(S_x(msg))$.
- Security assumption: Even though P_x is known publicly for all x, it is *hard* for an intruder to ascertain S_x from P_x. Only the owner x can compute S_x in a practical amount of time.
- Data Confidentiality: Assume A is the sender and B is the recipient of a message M. Then A encrypts by applying P_B of B, *i.e.* $C = P_B(M)$, Thus, only B can decode this message with S_B (*i.e.*, $S_B(P_B(M)) = M$)
- Digital signatures can also be implemented with Public-key cryptosystems: A can send a message M by encrypting it as $S_A(M)$. Note that any machine with P_A can decrypt this message. However, only A could have sent this message, since S_A is a secret known only to A.

Public-key Cryptosystems: RSA

A popular public-key cryptosystem is the Rivest–Shamir–Adleman algorithnm (authors given Turing Award in 2002)

1. Select two very large primes p and q [Use the probabilistic Miller-Rabin or Solovay-Strassen]
2. Compute $n = pq$. Compute $\phi(n) = (p-1)(q-1)$.
3. Choose an odd e s.t. $1 < e < \phi(n)$ and $gcd(e, \phi(n)) = 1$ [Use Euclid's gcd computation to select e]
4. Compute d as the multiplicative inverse of e, modulo $\phi(n)$. That is $ed \equiv 1 \ (mod\ \phi(n))$ [Apply Extended Euclid to solve for x s.t. $gcd(e, \phi(n)) = 1 = ex + \phi(n)y$]
5. Publish the public key $P = (e, n)$ of the participant
6. Publish the private key $S = (d, n)$ of the participant
7. The domain of a message $\mathcal{D}$ is $\mathbb{Z}_n = \{0, 1, 2, \cdots, n-1\}$.
8. Thus $P(M) = M^e \ (mod\ n) = C$. And $S(C) = C^d \ (mod\ n)$.

Why does RSA work?

- Note $P(S(M)) = S(P(M)) = M^{ed}\ (mod\ n)$.
- Also, $ed = 1 + k(p-1)(q-1)$
- So
  $M^{ed}\ (mod\ p) = M(M^{p-1})^{k(q-1)}\ (mod\ p) = M(1)^{k(q-1)}\ (mod\ p)$
  [Follows from Fermat's Theorem]
- Repeating the same argument, we will get
  $M^{ed}\ (mod\ q) = M\ (mod\ q)$. For all M

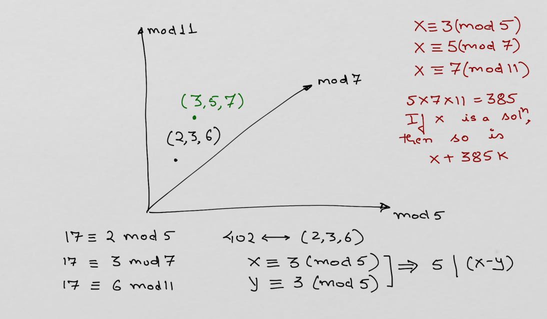
$$M^{ed} \equiv M\ (mod\ p)$$
$$M^{ed} \equiv M\ (mod\ q)$$

- From Chinese remainder theorem, $M^{ed} \equiv M\ (mod\ n)$

Chinese Remainder Theorem

If $p_1, p_2, \cdots, p_k$ are pairwise relatively prime, then for any integers $a_1, a_2, \cdots, a_k$, the set of equations: $x \equiv a_i \ mod \ p_i$ has a unique solution modulo $p_1 p_2 \cdots p_k$.

- Eg: $x \equiv 3 \ mod \ 5$
  $x \equiv 5 \ mod \ 7$
  $x \equiv 7 \ mod \ 11$



Chinese Remainder Theorem (Cont.)

- Consider three numbers $x_1, x_2, x3$ corresponding to the coordinates $(1,0,0), (0,1,0), (0,0,1)$ respectively.
- Then the point corresponding to the point $(3,5,7)$ is $3x_1 + 5x_2 + 7x_3$.
- For x_1:

$$x_1 \equiv 1 \ (mod\ 5) \quad (1)$$
$$x_1 \equiv 0 \ (mod\ 7) \quad (2)$$
$$x_1 \equiv 0 \ (mod\ 11) \quad (3)$$

- $7 * 11 \mid x_1$. Thus $77x_1' \equiv 1 \ (mod\ 5)$. Using eqn (1), we get $x_1 = 231$.
- Similarly, one can compute $x_2 = 330$ and $x_3 = 210$.
- Thus, $3x_1 + 5x_2 + 7x_3 = 3813$. Take factors of 385 out. The smallest positive number left is: 348 (solution to the original set of modular linear equations).

Chinese Remainder Theorem (Cont.)

- Provides a correspondence between a system of equations modulo a set of pairwise relative prime and an equation modulo the product of those pairwise relative primes
- "Structure Theorem" – describes the structure of $\mathbb{Z}_n$ is identical to that of $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$.
- As a result: Design of efficient algorithms (since working with $\mathbb{Z}_{n_i}$ is more efficient that working with $\mathbb{Z}_n$.

Security and Runtime Complexity of RSA

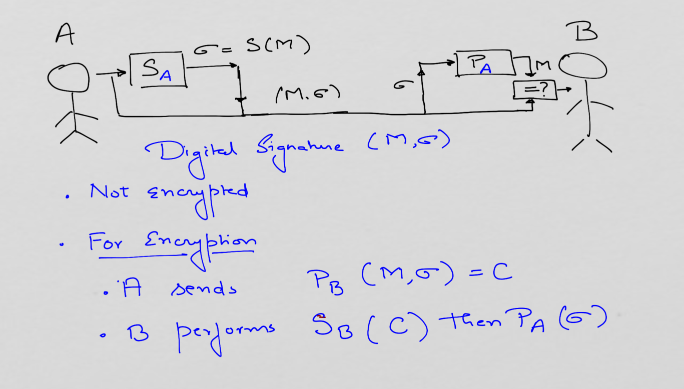
Security of RSA

- $M^{ed} \equiv M \ (mod \ n)$. To derive e and d, one will have to factor n. Typically, n is a product of two 1024 bit (300 digit) primes.

Runtime Complexity

- Applying P requires $O(1)$ modular multiplications. Applying S requires $O(\beta)$ modular multiplications (where β is the number of bits used to represent n).

Digitial Signatures



- A's digital signature for message M: $(M, S_A(M))$
- B upon receiving the signature decrypts $P_A(S_A(M))$ and performs the check $P_A(S_A(M)) \stackrel{?}{=} M$

Digital Signatures (continued)

- Note however, that the message M is sent over as plaintext
- An efficient approach is to combine data encryption with *Cryptographic hash functions*.
- CHF: allow fixed-length message fingerprints (provides *message integrity*
- A's digital signature for the message M: $\sigma = S_A(h(m))$. A sends the message $C = P_B(M, \sigma)$.
- Now, no eavesdropper can get the message in plaintext.
- Upon receiving the ciphertext, B decrypts by performing $S_B(C)$ and extracts the message: $(M, S_A(h(M)))$. It further performs the check $h(m) \stackrel{?}{=} P_A(S_A(h(m)))$.

Digital Certificates

- Certificates makes distributing public keys much easier
- An actor A can obtain a signed message from a publicly trusted authority T stating: A's public key is P_A.
- Actor A can include this certificate in her signed message.
- The recipient can now verify her signature with A's public key and the certificate from T.
- The recipient can now trust that A's key is indeed hers because of public trust in T.