Lecture 25: Primality Testing

12/3/2025

1 Primality Testing

Problem: Given an integer N, check if N is a prime number.

Note that the representation size is $\log N$ bits. So an efficient algorithm runs in $\operatorname{poly}(\log N)$ time.

- PRIMES \in P: "20134" interpreted as a binary integer is prime
- COMPOSITE: "20117" \in PRIMES

It is easy to see that COMPOSITE is in NP since one can prove that N is COMPOSITE by exhibiting x, y such that $N = x \cdot y$.

Not obvious that PRIMES is in NP. Vaughn Pratt in 1975 showed that PRIMES is in NP.

Question: Is there a poly time algorithm to check if N is a prime?

It had to wait till 2002 for a deterministic poly time algorithm due to Agarwal, Kayal, and Saxena

However, a randomized poly time algorithm was known since 1977 due to Solovay and Strassen and Miller and Rabin.

We need some number theoretic and group theoretic background.

2 Background

Claim 1. Given non-negative integers a, k, n, can compute $a^k \mod n$ efficiently.

Proof. Exercise. \Box

2.1 Euclid's Algorithm

Given integers a, b > 0, Euclid's algorithm can be used to obtain the following:

Theorem 1. Given integers a, b > 0, \exists integers x, y such that gcd(a, b) = ax + by. Moreover, x, y can be computed in poly time.

2.2 Groups

Let n > 0 be a positive integer. $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$ defines an additive abelian group under the operation mod n.

Let $\mathbb{Z}_n^* = \{a > 0 : a \in \mathbb{Z}_n, \gcd(a, n) = 1\}$ be the set of numbers that are relatively prime to n and in \mathbb{Z}_n .

Claim 2. \mathbb{Z}_n^* is a group under multiplication mod n.

Proof. Recall Euclid's gcd algorithm. Given a, b, it returns gcd(a, b) and can be implemented to run in poly time. As a byproduct, it also returns x, y such that ax + by = gcd(a, b).

Now consider $a \in \mathbb{Z}_n^*$ and b = n. $\exists x, y$ such that ax + yn = 1, thus $ax \equiv 1 \pmod{n}$. $x \mod n$ is a candidate for the inverse of a in \mathbb{Z}_n^* .

Cannot have x, x' such that $xa \equiv 1$ and $x'a \equiv 1 \pmod{n}$ because $(x - x')a \equiv 0 \pmod{n}$ but $\gcd(a, n) = 1 \Rightarrow x = x'$.

Corollary 1. $\mathbb{Z}_p^* = \{1, 2, \dots, p-1\}$ is a group. \mathbb{Z}_p is a field.

Remark 1. Proof also shows that given $a \in \mathbb{Z}_n^*$, one can find a^{-1} efficiently.

2.3 Euler's Totient Function

Definition 1 (Euler totient function). $\phi(m)$ for integer m > 0 is $|\mathbb{Z}_m^*|$.

Properties of ϕ :

- $\phi(1) = 1$
- For prime p: $\phi(p) = p 1$
- For prime p and k > 0: $\phi(p^k) = p^k p^{k-1}$
- For relatively prime n, m: $\phi(nm) = \phi(n)\phi(m)$

Via above properties:

Theorem 2. If n has prime factorization $p_1^{k_1}p_2^{k_2}\cdots p_t^{k_t}$, then

$$\phi(n) = n \prod_{i=1}^{t} \left(1 - \frac{1}{p_i} \right)$$

2.4 Lagrange's Theorem

Theorem 3 (Lagrange). Let G be a finite group and let $H \subseteq G$ be a subgroup. Then |H| divides |G|.

Definition 2. A group G is *cyclic* if \exists an element $g \in G$ such that $\forall a \in G$, \exists integer k such that $g^k = a$. g is called a *generator* for the group.

Definition 3. Given group G and $a \in G$, ord(a) is smallest integer k such that $g^k = 1$ (identity).

For any $a \in G$, $H(a) = \{1, a, a^2, \dots, a^{\operatorname{ord}(a)-1}\}$ forms a cyclic subgroup of G. Therefore we have: $\operatorname{ord}(a)$ divides |G| for $a \in G$.

This implies:

Theorem 4 (Euler). For any $a \in \mathbb{Z}_m^*$, $a^{\phi(m)} \equiv 1 \pmod{m}$.

Corollary 2 (Fermat). For any prime $p: a^{p-1} \equiv 1 \pmod{p}$ or $a^p \equiv a \pmod{p}$.

Another useful lemma:

Lemma 1. For any n > 0: $\sum_{d|n} \phi(d) = n$.

Proof. Consider integer $d \in \{1, 2, ..., n\}$. Let $A_d = \{1 \le x \le n : \gcd(x, n) = d\}$. $A_d = \emptyset$ if d is not a divisor of n. $\{A_d:d|n\}$ partition $\{1,2,\ldots,n\}$. Hence $\sum_d |A_d| = n$.

Not difficult to see $|A_d| = \phi(n/d)$. Hence

$$\sum_{d|n} \phi(d) = \sum_{d|n} \phi(n/d) = \sum_{d|n} |A_d| = n$$

Cyclic Groups 2.5

It is easy to see that \mathbb{Z}_n for any n is cyclic. However, \mathbb{Z}_n^* need not be cyclic in general. However:

Theorem 5. Let p be prime. Then \mathbb{Z}_p^* is cyclic.

Proof. Recall $|\mathbb{Z}_p^*| = p - 1$. For k|(p-1), let $O_k = \{a \in \mathbb{Z}_p^* : \operatorname{ord}(a) = k\}$ be the set of elements with order k.

From previous lemma: $\sum_{k|(p-1)} \phi(k) = p-1$. We will show later that $|O_k| = 0$ or $\phi(k)$.

$$\sum_{k} |O_k| = \sum_{k|(p-1)} \phi(k) = p - 1$$

and $|O_k| = \phi(k)$ for k|(p-1). Thus $|O_{p-1}| = \phi(p-1) \ge 1$, so $\exists g \in \mathbb{Z}_p^*$ with order p-1.

Claim 3. $|O_k| = 0 \text{ or } \phi(k)$.

All elements in O_k are roots of the polynomial $X^k \equiv 1 \pmod{p}$ over the field \mathbb{Z}_p . Suppose $|O_k| \ge 1$. Then $\exists a$ root for the above polynomial, and further all the roots are $\{1, a, a^2, \dots, a^{k-1}\}$. Since these are distinct, $\operatorname{ord}(a) = k$. Note that $a^i \in O_k$ if $\gcd(i,k) = 1$. Thus $|O_k| = \phi(k)$ if $|O_k| \neq 0.$

A number theoretic theorem:

Theorem 6. \mathbb{Z}_n^* is cyclic iff $n \in \{1, 2, 4, p^k, 2p^k\}$ for integer k and odd prime p.

2.6 Chinese Remainder Theorem

Theorem 7 (Chinese Remainder Theorem). In a ring where $n = n_1 n_2 \cdots n_k$ are pairwise coprime, i.e., $gcd(n_i, n_j) = 1$ for $i \neq j$. For any sequence r_1, \ldots, r_k where $r_i \in \mathbb{Z}_{n_i}$, there exists a unique $r \in \mathbb{Z}_n$ s.t. $r \equiv r_i \pmod{n_i}$ for i = 1 to k. Moreover, given $\{r_i\}_i$, r can be computed efficiently.

Proof. First we consider showing one r exists. Since n/n_i is coprime to n_i , a multiplicative inverse m_i in \mathbb{Z}_{n_i} to n/n_i exists. Let m_i be that inverse: $(n/n_i) \cdot m_i \equiv 1 \pmod{n_i}$.

Also $(n/n_i) \cdot m_i \equiv 0 \pmod{n_j}$ for $j \neq i$.

Thus $r = \sum_{i} r_i m_i (n/n_i) \mod n$ satisfies the desired congruences.

To see uniqueness, we do a counting argument. How many distinct (r_1, \ldots, r_k) are there? $n_1 \cdot n_2 \cdots n_k$. For each $r \in \mathbb{Z}_n$, but there are only n elements in \mathbb{Z}_n , and for a given r we have only one $(r_1, ..., r_k)$.

$$\mathbb{Z}_n$$
 is isomorphic to $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$

2.7 Quadratic Residues

Definition 4. A residue $a \in \mathbb{Z}_m^*$ is a *quadratic residue* if \exists a number x such that $x^2 \equiv a \pmod{m}$. In other words, a is a quadratic residue if it has a square root.

Lemma 2. Let p be a prime and consider generator g for \mathbb{Z}_p^* . Then g^k is a quadratic residue iff k is even.

Proof. It is easy to see that if k is even then $g^k = (g^{k/2})^2$ is a square root of g^k . Suppose k is odd. If g^k is a quadratic residue then $g^k = (a)^2 = g^{2h}$ for some h since g is a generator. Thus k = 2h, contradiction.

Corollary 3 (Euler). For $a \in \mathbb{Z}_p^*$, a is a quadratic residue iff $a^{(p-1)/2} \equiv 1 \pmod{p}$.

Proof. $a=g^k$ where g is a generator, so $g^{p-1}=1$ gives $g^{(p-1)/2}=\pm 1$. For a generator g, $g^{(p-1)/2}\equiv -1 \pmod p$.

Definition 5 (Legendre symbol). For prime p and $a \in \mathbb{Z}_p^*$ we define $\left(\frac{a}{p}\right)$ where

$$\begin{pmatrix} \frac{a}{p} \end{pmatrix} = \begin{cases} 1 & \text{if } a \text{ is quadratic residue} \\ -1 & \text{if } a \text{ is not a quadratic residue} \end{cases}$$

Equivalently, $\binom{a}{p} \equiv a^{(p-1)/2} \pmod{p}$ where we interpret p-1 as -1.

Now we consider a not necessarily prime odd number.

Definition 6 (Jacobi symbol). Let n be an odd number with prime factorization $n = p_1^{k_1} p_2^{k_2} \cdots p_t^{k_t}$. For $a \in \mathbb{Z}_n^*$:

$$\left(\frac{a}{n}\right) = \prod_{i=1}^{t} \left(\frac{a}{p_i}\right)^{k_i}$$

Note that $\left(\frac{a}{n}\right)$ is the same as the Legendre symbol when n is an odd prime. It is also ± 1 .

Even though the definition of the Jacobi symbol involves the prime factorization, given a, n one can compute $\left(\frac{a}{n}\right)$ in poly time.

Theorem 8. Given a, n where n is odd and gcd(a, n) = 1, there is a poly time alg to compute $\left(\frac{a}{n}\right)$.

One can derive the above from properties of the Jacobi symbol (see Motwani, Raghavan).

Definition 7. For an odd number n, define

$$J_n = \left\{ a \in \mathbb{Z}_n^* : \left(\frac{a}{n}\right) = 1 \right\}$$

Note that $J_n = \mathbb{Z}_n^*$ when n is prime.

A key observation is:

Lemma 3. J_n is a subgroup of \mathbb{Z}_n^* and is a proper subgroup if n is composite. Hence if n is composite, $|J_n| \leq |\mathbb{Z}_n^*|/2$.

3 Solovay-Strassen Algorithm

The preceding observation leads to first RP algorithm for COMPOSITE due to Solovay and Strassen.

Algorithm: Input $n \geq 2$

- 1. If n is even, output Composite
- 2. Pick a random $a \in \{2, \ldots, n-1\}$
- 3. If $gcd(a, n) \neq 1$, output Composite
- 4. Compute $a^{(n-1)/2} \mod n$
- 5. Compute $\left(\frac{a}{n}\right)$
- 6. If $a^{(n-1)/2} \not\equiv \left(\frac{a}{n}\right) \pmod{n}$, then output Composite
- 7. Else output Prime

It is clear that algorithm outputs Prime for a prime, but for a composite it will err with prob $\leq 1/2$.

Theorem 9. $COMPOSITE \in RP$.

By repeating we can reduce the error.

4 Miller-Rabin Test

We will assume n is odd and > 5.

Simplest randomized test is to pick a number $a \in \{2, ..., n-1\}$ and check if $gcd(a, n) \neq 1$. Test will succeed with prob $\geq 1 - \phi(n)/n$ but we can have $\phi(n) \approx n$. For instance, n = pq where p, q are prime, then $\phi(n) = (p-1)(q-1)$. So we need some property.

4.1 Fermat Test

Pick $a \in \{2, \ldots, n-1\}$ randomly. If $a^{n-1} \not\equiv 1 \pmod n$, output Composite; else Prime.

If n is prime then correctly says it is prime, and if it says it is composite then it is correct, but it may say prime even when n is composite.

What is the probability?

Let
$$F_n = \{a \in \mathbb{Z}_n^* : a^{n-1} \equiv 1 \pmod{n}\}.$$

Claim 4. F_n is a subgroup of \mathbb{Z}_n^* .

Proof. If
$$a, b \in F_n$$
, then $ab \in F_n$. Also a^{-1} is inverse of a and $a^{-1} \in F_n$.

Suppose F_n is a proper subgroup of \mathbb{Z}_n^* . Then $|F_n| \leq |\mathbb{Z}_n^*|/2$ by Lagrange's theorem, and algorithm will have constant probability of success.

But what if $F_n = \mathbb{Z}_n^*$? Are there any such composite n?

Yes, they are called *Carmichael numbers*. There are infinitely many. Smallest is $561 = 3 \cdot 11 \cdot 17$.

Thus we need a test that handles Carmichael numbers.

Another property of primes is that \mathbb{Z}_p is a field, and $x^2 \equiv 1 \pmod{p}$ has only two roots: x = 1 and $x = p - 1 \equiv -1 \pmod{p}$.

If n is composite then there can be non-trivial square roots. Ex: n = 91, then 1, 27, 64, 90 are all square roots of 1.

If we find a non-trivial square root of 1 mod n, then n is composite. If n = pq, then by CRT only 4 non-trivial square roots.

4.2 Euler Test

Pick $a \in \{2, \ldots, n-1\}$. If $a^{(n-1)/2} \not\equiv \pm 1 \pmod{n}$, output Prime; else Composite.

At least as good as Fermat test. But why only $a^{(n-1)/2}$ and not $a^{(n-1)/4}$? First may not be even, but if $(n-1)/2^i$ is odd then we can try.

4.3 Rabin-Miller Test

n is odd. Write $n-1=2^u \cdot k$ where k is odd.

Algorithm:

- 1. Pick $a \in \{2, \dots, n-1\}$
- 2. If $gcd(a, n) \neq 1$, output Composite
- 3. $b_0 = a^k \mod n$
- 4. If $b_0 = 1$, output Prime
- 5. For i = 1 to u 1 do:
 - $b_i = b_{i-1}^2 \mod n$
 - If $b_i = -1$, output Prime
 - If $b_i = 1$, return Composite
- 6. Return Composite since either $b_u \not\equiv 1 \pmod{n}$ or $b_u = 1$ and $b_{u-1} \not\equiv -1$

Theorem 10. If n is prime, alg outputs Prime with prob 1. If n is composite, alg outputs Prime with $prob \leq 1/4$.

Sketch of the analysis:

Fix $a \in \mathbb{Z}_n^*$. Let $b_0 = a^k \mod n$, $b_1 = a^{k \cdot 2} = a^{2k} \mod n$, $b_2 = a^{k \cdot 2^2} = a^{4k} \mod n$, ..., $b_u = a^{k \cdot 2^u} = a^{n-1} \mod n$.

If b_i becomes 1, then b_{i+1}, \ldots are all 1. When do we find a non-trivial square root of 1? If $b_i = 1$ and $b_{i-1} \neq \pm 1$. This is precisely where the alg outputs Composite. Also if $b_u = a^{n-1} \not\equiv 1 \pmod{n}$. Example: $n = 325 = 5^2 \cdot 13$, $n - 1 = 324 = 81 \cdot 2^2$.

Definition 8. Let $n \geq 3$ be an odd composite. Let $n-1=2^k \cdot q$ where q is odd and $k \geq 1$. A number $a \in \mathbb{Z}_n^*$ is an RM witness for n if $a^q \not\equiv 1 \pmod n$ and $a^{2^iq} \not\equiv -1 \pmod n$ for $0 \leq i < k$.

If n is composite and a is not an RM witness for n, then a is an RM liar for n.

Lemma 4. If a is an RM witness for n, then n is composite.

4.4 Analysis

We will prove a weaker theorem:

Theorem 11. Let n > 3 be odd composite, and let L(n) be the set of RM liars for n. Then $|L(n)| \leq |\mathbb{Z}_n^*|/4$.

The difficulty is that L(n) is not a subgroup of \mathbb{Z}_n^* . Thus to prove the theorem we identify a proper subgroup S of \mathbb{Z}_n^* and argue that $L(n) \subseteq S$.

We consider two cases:

Case 1: n is not a Carmichael number. If a is a RM liar then it is also a Fermat liar, and we argued that $|F_n| \leq |\mathbb{Z}_n^*|/2$ when n is not a Carmichael number.

Case 2: n is a Carmichael number.

However, this requires us to understand Carmichael numbers. We will state and not prove:

Theorem 12. Suppose $n = p_1^{k_1} p_2^{k_2} \cdots p_t^{k_t}$ where each p_i is an odd prime.

- (a) n is Carmichael iff $\phi(p_i^{k_i})|(n-1)$ for i=1 to t.
- (b) If n is Carmichael, then $n = p_1 p_2 \cdots p_t$ and $(p_i 1)|(n 1)$ for i = 1 to t. In particular, $t \ge 3$.

Our goal is to find a proper subgroup S of \mathbb{Z}_n^* s.t. $L(n) \subseteq S$.

Let i_0 be maximal i > 0 such that there is some RM liar a_0 with $a_0^{2^{i_0}q} \equiv -1 \pmod{n}$. Since n is odd composite, $i_0 \geq 1$, so i_0 exists.

Since n is a Carmichael number, $a^{n-1} \equiv 1 \pmod{n}$ and hence $0 \le i_0 \le k$.

Let $B_n = \{ a \in \mathbb{Z}_n^* : a^{2^{i_0}q} \equiv \pm 1 \pmod{n} \}.$

Lemma 5.

- (i) $L(n) \subseteq B_n$
- (ii) B_n is a subgroup of \mathbb{Z}_n^*
- (iii) $\mathbb{Z}_n^* \setminus B_n \neq \emptyset$

Proof. (i) Let $a \in L(n)$. Case 1: $a^q \equiv 1 \pmod{n}$. Then $a^{2^{i_0}q} \equiv 1 \pmod{n}$, hence $a \in B_n$. Case 2: $a^{2^{i_q}} \equiv -1 \pmod{n}$ for some i. By defin of i_0 : $0 \le i \le i_0$. If $i = i_0$ then $a \in B_n$. If $i < i_0$, then $a^{2^{i_0}q} \equiv 1 \pmod{n}$ and hence $a \in B_n$.

- (ii) B_n is a subgroup: $1 \in B_n$ trivially. If $a, b \in B_n$, then $ab \in B_n$ (easy).
- (iii) We know that n has at least 3 prime factors, hence $n = n_1 n_2$ where n_1, n_2 are odd and $gcd(n_1, n_2) = 1$. Recall a_0 is a RM liar with $a_0^{2^{i_0}q} \equiv -1 \pmod{n}$. Let $\bar{a} = a_0 \mod n_1$. By CRT, \exists unique $a \in \mathbb{Z}_n$ s.t. $a \equiv \bar{a} \pmod{n_1}$ and $a \equiv 1 \pmod{n_2}$.

We claim $a \in \mathbb{Z}_n^* \setminus B_n$.

Computing mod n_1 : since $\bar{a} = a_0 \mod n_1$, $a^{2^{i_0}q} \equiv -1 \pmod {n_1}$.

Computing mod n_2 : since $a \equiv 1 \pmod{n_2}$, $a^{2^{i_0}q} \equiv 1 \pmod{n_2}$.

Thus $a^{2^{i_0}q} \equiv -1 \pmod{n_1}$ and $a^{2^{i_0}q} \equiv 1 \pmod{n_2}$, so $a \notin B_n$.

Further, $a^{n-1} \equiv 1 \pmod{n_1}$ and $a^{n-1} \equiv 1 \pmod{n_2}$. By CRT, $a^{n-1} \equiv 1 \pmod{n}$, so $\gcd(a,n)=1$ and $a \in \mathbb{Z}_n^*$.

Alternate Proof (Keith Conrad)

The preceding proof used characterization of Carmichael numbers and is from Dietzfelbinger's book. We give an alternate proof from the notes of Keith Conrad that avoids the use of Carmichael numbers.

We again consider 2 cases:

Case 1: $n = p^k$ for $k \ge 2$, p prime power. We claim n is not Carmichael; we can use the fact that $|F_n| \leq |\mathbb{Z}_n^*|/2$.

To see this, it suffices to exhibit some $a \in \mathbb{Z}_n^*$ such that $a^{n-1} \not\equiv 1 \pmod{n}$.

Consider $(1+p)^{n-1}$. By Binomial expansion:

$$(1+p)^{n-1} = 1 + (n-1)p + {n-1 \choose 2}p^2 + \cdots$$

Hence $(1+p)^{n-1} \equiv 1 + (n-1)p \pmod{p^2}$. If k = 2: $(1+p)^{n-1} \equiv 1 + (n-1)p \not\equiv 1 \pmod{p^2}$. If k > 2: $(1+p)^{n-1} \equiv 1 + (n-1)p \pmod{p^k}$ since $(n-1)p = p^k - p$ and $p^k|p^k$ but $p^k \nmid p^{k-1}$.

Case 2: n is not a prime power. Write $n = p^s n_2$ where p does not divide n_2 , so $n = n_1 n_2$ where n_1, n_2 are odd and n_1, n_2 relatively prime.

Let i_0 be maximal in $\{0,\ldots,u-1\}$ such that there is some $a\in\mathbb{Z}_n^*$ such that $a^{2^{i_0}q}\equiv -1$ (mod n). Since $(-1)^2 = 1 \mod n$, i_0 exists and $a \in \mathbb{Z}_n^*$. Let $B_n = \{a \in \mathbb{Z}_n^* : a^{2^{i_0}q} \equiv \pm 1 \pmod n\}$.

Lemma 6.

- (i) $L(n) \subset B_n$
- (ii) B_n is a subgroup of \mathbb{Z}_n^*
- (iii) $\mathbb{Z}_n^* \setminus B_n \neq \emptyset$

Proof. (i) Let $a \in L(n)$. Case 1: $a^q \equiv 1 \pmod{n}$. Then $a^{2^{i_0}q} \equiv 1 \pmod{n}$, hence $a \in B_n$. Case 2: $a^{2^iq} \equiv -1 \pmod{n}$ for some i. By defin of i_0 : $0 \le i \le i_0$. If $i = i_0$ then $a \in B_n$. If $i < i_0$, then $a^{2^{i_0}q} \equiv 1 \pmod{n}$ and hence $a \in B_n$.

- (ii) B_n is a subgroup: $1 \in B_n$ trivially. If $a, b \in B_n$, then $ab \in B_n$ (easy).
- (iii) $n = n_1 n_2$ where n_1, n_2 are odd and $gcd(n_1, n_2) = 1$. Recall a_0 is a RM liar with $a_0^{2^{i_0}q} \equiv -1$ (mod n). Let $\bar{a} = a_0 \mod n_1$. By CRT, \exists unique $a \in \mathbb{Z}_n$ s.t. $a \equiv \bar{a} \pmod{n_1}$ and $a \equiv 1 \pmod{n_2}$. We claim $a \in \mathbb{Z}_n^* \setminus B_n$.

Computing mod n_1 : since $\bar{a} = a_0 \mod n_1$, $a^{2^{i_0}q} \equiv -1 \pmod {n_1}$.

Computing mod n_2 : since $a \equiv 1 \pmod{n_2}$, $a^{2^{i_0}q} \equiv 1^{2^{i_0}q} \equiv 1 \pmod{n_2}$.

Thus $a^{2^{i_0}q} \equiv -1 \pmod{n_1}$ and $a^{2^{i_0}q} \equiv 1 \pmod{n_2}$, so $a \notin B_n$.

Further, $a^{n-1} \equiv a_0^{n-1} \pmod{n_1}$ and $a^{n-1} \equiv 1 \pmod{n_2}$. By CRT, $a^{n-1} \equiv 1 \pmod{n}$ (need to verify $a_0^{n-1} \equiv 1 \pmod{n_1}$, which follows from Euler's theorem), so $\gcd(a,n) = 1$ and $a \in \mathbb{Z}_n^*$.