Generation of Prime Numbers

Prof. Ashok K Bhateja, IIT Delhi

Generation of large primes & primality test

- The sieve of Eratosthenes
- Trial Division test
- Fermat’s primality test
- Solovay-Strassen test
- Miller-Rabin test

The sieve of Eratosthenes

1. Create a list of consecutive integers from 2 to n i.e., (2, 3, 4, …, n).
2. Initially, let p equal 2, the first prime number.
3. Starting from p^2, count up in increments of p and mark each of these numbers greater than or equal to p^2 itself in the list.
4. Find the first number greater than p in the list that is not marked.
   - If there was no such number, stop.
   - Otherwise, let p now equal this number (which is the next prime), and repeat from step 3.
5. the numbers remaining not marked in the list are all the primes below n.

The sieve of Eratosthenes for n = 20

	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
11	12	13	14	15	16	17	18	19	20

	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
11	12	13	14	15	16	17	18	19	20

	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
11	**12**	13	14	15	16	17	**18**	19	20

	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
11	**12**	13	14	15	16	17	**18**	19	20

The primes are: 2, 3, 5, 7, 11, 13, 17, 19

Trial Division Test

- If n is not prime, then at least one of the factors of n is at most as large as $\sqrt{n}$.
- Divide the candidate number by only the primes up to its square root.
- In the worst case, trial division is a laborious algorithm. For an n-bit number a, if it starts from two and works up only to the square root of a, the algorithm requires

$$\pi\left(2^{n/2}\right) \approx \frac{2^{n/2}}{(n/2)\ln 2}$$

Trial Division Test

To check Integer $n \geq 2$ is prime

$i \leftarrow 2$

while $i.i \leq n$ do

    if i divides n then

        return COMPOSITE

    end if

    $i \leftarrow i + 1$

end while

return PRIME

Probabilistic primality tests

- Probable prime
  - believed to be prime based on a probabilistic primality test.
  - an integer that satisfies a specific condition that is satisfied by all prime numbers, but which is not satisfied by most composite numbers.
- Witnesses to the compositeness of *n*
  - Let n be an odd composite integer. An integer a, coprime to n, is Fermat witness of n, if the probabilistic test outputs composite.
  - Let $n = 1387$. Since $2^{1386} \equiv 1 \bmod 1387$, implies n may be prime. However, $3^{1386} \equiv 875 \neq 1 \bmod 1387$, so 1387 is composite with 3 as a Fermat witness.

Algorithm: Fermat primality testing

for i = 1 to t

choose a random integer a, $2 \le a \le n - 1$.

compute $r \equiv a^{(n-1)} \bmod n$

if $r \neq 1$ then return ("composite")

return("prime")

- If n is prime, then the Fermat primality test always outputs prime. If n is composite, then the algorithm outputs prime with probability at most 1/2.

Fermat's Test : When will it give error?

- If the number is prime the algorithm will always give the output as "PRIME".
- If the input number is composite, the algorithm might claim that the number is prime. [give an error]
- Why is this error generated? Due to the presence of F-Liars
- For an odd composite number n, an element a, $1 \leq a \leq n - 1$, is F-liar if $a^{(n-1)} \bmod n \equiv 1$ and n is called Fermat pseudoprime to base a.
- Example: $n = 341$ $(= 11 \times 31)$ is a pseudoprime to the base 2 since $2^{340} \equiv 1 \pmod{341}$.

Fermat’s Test : Error Probability

- Theorem: If a composite integer $n > 1$ has a Fermat witness that is relatively prime to n then the proportion of integers from 2 to $n - 1$ that are Fermat witnesses for n is over 50%.
- If over half the integers in $\{2, . . . , n - 1\}$ are Fermat witnesses for n, then the probability of not finding a Fermat witness among, say, k random choices is smaller than $(½)^k$.
- So, we might say that n appears to be prime with “probability” at least $1 - (½)^k$. For k =10, it is ≈ 0·99902.

Carmichael function

- Let n be a positive integer. The Carmichael function $\lambda(n)$ is the least positive integer m such that $a^m \equiv 1(\text{mod } n)$ for all integers a coprime to n.

  i.e., $a^{\lambda(n)} \equiv 1(\text{mod } n)\ \forall\ a$ coprime to n.

- $\lambda(8) = 2$, because for any number a coprime to 8 it holds that $a^2 \equiv 1 \pmod 8$.

  $1^2 \equiv 1 \pmod 8$, $3^2 = 9 \equiv 1 \pmod 8$, $5^2 = 25 \equiv 1 \pmod 8$ and $7^2 = 49 \equiv 1 \pmod 8$.

- $\varphi(8) = 4$, because there are 4 numbers less than and coprime to 8 i.e., 1, 3, 5, and 7.

- Euler's theorem assures that $a^4 \equiv 1 \pmod 8$ for all a coprime to 8, but 4 is not the smallest such exponent.

Computing $\lambda(n)$

Any $n > 1$ can be written as $n = p_1^{\alpha_1}.p_2^{\alpha_2} \dots p_k^{\alpha_k}$ be the prime factorization of n. Then

$$\lambda(n) = lcm\left\{\lambda\left(p_1^{\alpha_1}\right), \lambda(p_2^{\alpha_2}), \dots, \lambda\left(p_k^{\alpha_k}\right)\right\} \text{ where } n = \prod_{i=1}^{k} p_i^{\alpha_i}$$

$$\lambda(p^\alpha) = \begin{cases} \varphi(p^\alpha) & \text{if } \alpha \le 2 \text{ or } p \ge 3 \\ \frac{1}{2}\varphi(p^\alpha) & \text{if } p = 2 \text{ and } \alpha \ge 3 \end{cases}$$

$$\lambda(mn) = lcm\ (\lambda(m), \lambda(n))$$

Example: $\lambda(360) = \text{lcm}\ (\lambda(2^3), \lambda(3^2), \lambda(5)) = \text{lcm}(2, 6, 4) = 12.$

$\lambda(561) = \text{lcm}\ (\lambda(3), \lambda(11), \lambda(17)) = \text{lcm}(2, 10, 16) = 80$

Carmichael function

Theorem: If $\lambda(n) \mid (n-1)$, then $a^{n-1} \equiv 1 (\bmod n)$ for all a coprime to n.

Proof: $\lambda(n) \mid (n-1) \Rightarrow \lambda(n)\, k = (n-1)$

Therefore, $a^{n-1} = \left(a^{\lambda(n)}\right)^k \Rightarrow a^{n-1} \equiv 1 \bmod n$

i.e., If $\lambda(n) \mid (n-1)$ then Fermat's condition for prime is true whether n is prime or not.

Consider $n = 561$, $\lambda(561) = 80$, which divides 560.

$a^{560} \equiv 1 (\bmod 561)$ for all a coprime to 561.

But $561 = 3 \times 11 \times 17$ (not a prime)

Carmichael number

- Definition: A composite number n, which satisfies the relation $a^{(n-1)} \equiv 1 \bmod n$ for all integers a satisfying $\gcd(a, n) = 1$ with $1 < a < n$.
- The converse of Fermat's little theorem is not generally true, as it fails for Carmichael numbers.
- The first few Carmichael numbers are 561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341, ...
- Number of Carmichael numbers C(n) for sufficiently large n, is

$C(n) > n^{2/7}$ (Alford et al. 1994)

$C(n) < n \exp\left(-\frac{\ln n \ln \ln \ln n}{\ln \ln n}\right)$ (R.G.E Pinch)

Legendre symbol

- Let p be an odd prime and a is an integer. The Legendre symbol is defined as

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p|a \\ 1 & \text{if } a \in Q_p \\ -1 & \text{if } a \in \overline{Q}_p \end{cases}$$

Legendre symbol

Fact: Let p be an odd prime and a, $b \in$ Z. Then

(i) $\left(\frac{a}{p}\right) = a^{(p-1)/2} \bmod p$; $\left(\frac{1}{p}\right) = 1$, $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$

(ii) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$; if $a \in Z_n^*$, then $\left(\frac{a^2}{p}\right) = 1$.

(iii) If $a \equiv b \pmod p$, then $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$

(iv) $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$

(v) Law of quadratic reciprocity: If q is an odd prime distinct from p, then

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}$$

Jacobi Symbol

- Jacobi symbol is generalization of Legendre symbol .
- Definition Let $n \geq 2$ be odd integer and $n = p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}$ then Jacobi symbol of a & n is

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{e_1} \left(\frac{a}{p_2}\right)^{e_2} \cdots \left(\frac{a}{p_k}\right)^{e_k}$$

  If n is prime, then the Jacobi symbol is just the Legendre symbol.
- If m is composite and the Jacobi symbol (a/m) = -1, then a is quadratic non-residue modulo m.
- If a is quadratic residue modulo m & gcd (a, m) = 1, then (a/m) = 1,

  but if (a/m) = 1 then a may be quadratic residue or non-residue modulo m.
- Example: (2/15) = 1 and (4/15) = 1, but 2 N 15 and 4 R 15.

Properties of Jacobi symbol

1. $(a/n) = (b/n)$ if $a = b$ mod n.
2. $(1/n) = 1$ and $(0/n) = 0$.
3. $(2m/n) = (m/n)$ if $n = \pm 1$ mod 8.

   $(2m/n) = -(m/n)$ otherwise
4. (Quadratic reciprocity) If m and n are both odd, then

   $(m/n) = -(n/m)$ if both m and n are congruent to 3 mod 4

   $(m/n) = (n/m)$ otherwise.

Example: Compute Jacobi symbol (158/235)

$$\left(\frac{158}{235}\right) = -\left(\frac{79}{235}\right) \because n \neq \pm 1 \bmod 8$$

$$= \left(\frac{235}{79}\right) \because \text{both } m \,\&\, n \text{ are congruent to } 3 \bmod 4$$

$$= \left(\frac{10}{79}\right) \because 235 \equiv 10 \bmod 7\,9$$

$$= -\left(\frac{5}{79}\right) = -\left(\frac{79}{5}\right)$$

$$= -\left(\frac{4}{5}\right) = -\left(\frac{1}{5}\right) = -1$$

Solovay-Strassen test

- Fact (Euler's criterion) Let n be an odd prime.

  Then $a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \pmod{n}$

  for all integers, a which satisfy gcd $(a, n) = 1$.

- If gcd $(a, n) = 1$ and $a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \pmod{n}$ then n is said to be a Euler pseudoprime to the base a.

Algorithm Solovay-Strassen probabilistic primality test

INPUT: an odd integer $n > 3$ and security parameter $t \geq 1$.

for i from 1 to t

choose a random integer a, $2 \leq a \leq n - 2$

find gcd (a, n)

if gcd $(a, n) > 1$ then return ("composite")

compute $r = a^{(n-1)/2} \bmod n$

if $r \neq 1$ and $r \neq n - 1$ then return ("composite")

compute the Jacobi symbol $s = (a/n)$

if $r \neq s \pmod{n}$ then return ("composite")

return("prime")

Solovay-Strassen error-probability bound

- Fact: Let n be an odd composite integer. Then at most $\varphi(n)/2$ of all the numbers a, $1 \le a \le n - 1$, are Euler liars for n.
- Fact: Let n be an odd composite integer. The probability that Solovay-Strassen algorithm declares n to be "prime", with t bases, is less than $(1/2)^t$.
- Example: (Euler pseudoprime) The composite integer 91 (= 7 × 13) is a Euler pseudoprime to the base 9

  since $9^{45} = 1 \pmod{91}$ and $\left(\frac{9}{91}\right) = 1$.

Complexity of the Solovay-Strassen test

- GCD of two numbers can be calculated using the Euclidean algorithm having a complexity of $O(\log^2 n)$.
- Computing Jacobi symbol has the same complexity as the Euclidean algorithm.
- Multiplication of two numbers is always done modulo n and it takes $O(\log^2 n)$ time.
- For any a, we can compute $a^n \bmod n$ in $O(\log n)$ multiplications, by repeated squaring.
- Thus, this method of modular exponentiation can be done in $O(\log n \times \log^2 n) = \log^3 n$ for each value of a.
- The overall time-complexity of the Miller-Rabin algorithm is $O\,(t \cdot \log^3 n)$, t being the number of bases.

Miller-Rabin test

- It is a strong pseudoprime probabilistic test
- Fact: Let n be an odd prime and let $n - 1 = 2^s r$ where r is odd. Let a be any integer s.t. gcd $(a, n) = 1$. Then either
  $a^r \equiv 1 \pmod{n}$ or $a^{2^j \cdot r} \equiv -1 \bmod n$ for some j, $0 \leq j \leq s - 1$.

Def: Let n be an odd composite integer and let $n - 1 = 2^s r$ where r is odd. Let a be any integer in $[1, n - 1]$

i. If $a^r \not\equiv 1 \pmod{n}$ & $a^{2^j \cdot r} \not\equiv -1 \bmod n$ $\forall j$, $0 \leq j \leq s - 1$. then a is said a strong witness (to compositeness) for n.

ii. If $a^r \equiv 1 \pmod{n}$ or $a^{2^j \cdot r} \equiv -1 \bmod n$ for some j, $0 \leq j \leq s - 1$. then n is said to be a strong pseudoprime to the base a (i.e., n acts like a prime). The integer a is called a strong liar for n.

Number of Strong liars

- Fact: If n is an odd composite integer, then at most 1/4 of all the numbers a, $1 \leq a \leq n - 1$, are strong liars for n. In fact, the number of strong liars for n is at most $\varphi(n)/4$.
- Ex: Consider the composite integer $n = 91$ $(= 7 \times 13)$.

  $91 - 1 = 90 = 2 \times 45$, $s = 1$ and $r = 45$.

  Let $a = 9$, $9^r = 9^{45} \equiv 1 \pmod{91}$

  Implies 91 is a strong pseudoprime to the base 9.

  The set of all strong liars for 91 is:

  {1, 9, 10, 12, 16, 17, 22, 29, 38, 53, 62, 69, 74, 75, 79, 81, 82, 90}.
- The number of strong liars for 91 is $18 = \varphi(91)/4$.

Algorithm: Miller-Rabin probabilistic primality test

INPUT: An odd integer $n > 2$ and security parameter $t \geq 1$

write $n - 1 = 2^s r$ such that r is odd.

for i from 1 to t

  choose a random integer a, $2 \leq a \leq n - 2$

  compute $y = a^r \bmod n$

  if $y \neq 1$ and $y \neq n - 1$

    $j = 1$.

   while $j \leq s - 1$ and $y \neq n - 1$

    compute $y = y^2 \bmod n$

    if $y = 1$ then return (“composite”)

    $j = j + 1$

   if $y \neq n - 1$ then return (“composite”)

return(“prime”)

Miller-Rabin error-probability bound

- For any odd composite integer n, the probability that Miller Rabin primality test algorithm declares n to be “prime” is less than $(1/4)^t$
- For most composite integers n, the number of strong liars for n is actually much smaller than the upper bound of $\varphi(n)/4$.

  Consequently, the Miller-Rabin error-probability bound is much smaller than $(1/4)^t$ for most positive integers n.
- Example: (some composite integers have very few strong liars) The only strong liars for the composite integer $n = 105$ $(= 3 \times 5 \times 7)$ are 1 and 104. More generally, if $k \geq 2$ and n is the product of the first k odd primes, there are only 2 strong liars for n, namely 1 and $n - 1$.

Time complexity

- Multiplication of two numbers is always done modulo n and it takes $O(\log^2 n)$ time.
- For any a, we can compute $a^n \pmod{n}$ in $O(\log n)$ multiplications (modular exponentiation).
- Thus, this method of modular exponentiation can be done in $O(\log n \times \log^2 n) = \log^3 n$ for each value of a.
- The overall time-complexity of the Miller-Rabin algorithm is $O(t \cdot \log^3 n)$, t being the number of bases.

Comparison: Fermat, Solovay-Strassen and Miller-Rabin

- Fact: Let n be an odd composite integer.

  (i) If a is a Euler liar for n, then it is also a Fermat liar for n.

  (ii) If a is a strong liar for n, then it is also a Euler liar for n.

- Ex: For composite integer $n = 65$ $(= 5 \times 13)$, the Fermat liars for 65 are $\{1, 8, 12, 14, 18, 21, 27, 31, 34, 38, 44, 47, 51, 53, 57, 64\}$.

  The Euler liars for 65 are $\{1, 8, 14, 18, 47, 51, 57, 64\}$,

  while the strong liars for 65 are

  $\{1, 8, 18, 47, 57, 64\}$

Strong Prime

- Definition. A prime number *p* is said to be a strong prime if integers *r*, *s*, and *t* exist such that the following three conditions are satisfied:
  - p - 1 has a large prime factor, denoted r
  - p + 1 has a large prime factor, denoted s and
  - r - 1 has a large prime factor, denoted t
- A strong prime is a prime number that is greater than the arithmetic mean of nearest prime numbers i.e., next and previous prime numbers.
- The first few strong primes are

  11, 17, 29, 37, 41, 59, 67, 71, 79, 97, 101

Generation of Strong primes

Gordon's algorithm for generating a strong prime p

1. Generate two large random primes s and t of roughly equal bitlength.
2. Select an integer i_0.

   Find the first prime in the sequence $2it + 1$, for $i = i_0, i_0 + 1, i_0 + 2, \ldots$ Denote this prime by $r = 2it + 1$.
3. Compute $p_0 = (2s^{r-2} \bmod r)\, s - 1$.
4. Select an integer j_0

   Find the first prime in the sequence $p_0 + 2jrs$, for $j = j_0, j_0 + 1, j_0 + 2, \ldots$ Denote this prime by $p = p_0 + 2jrs$.
5. Return(p).