Computer Science & Engineering Department
IIT Kharagpur
Computational Number Theory: CS60094
Lecture VI

Instructor: Goutam Biswas *Spring Semester 2014-2015*

1 Test for prime I

Testing whether a natural number is prime is an old problem. Even today a deterministic testing takes a long time for an integer of large size ([LKP]). The first deterministic polynomial time algorithm for testing prime (AKS cyclotomic primality test) was proposed by Manindra Agrawal, Neeraj Kayal, and Nitin Saxena of IIT Kanpur in 2004.

1.1 Trial Division

We start with the old and well known simple algorithm of trial division.

```
isprimeTD($n$)
    $i \leftarrow 2$
    $sqrtN \leftarrow \lfloor\sqrt{n}\rfloor$
    while $i \leq sqrtN$
        if $n \bmod i = 0$ return 0
        $i \leftarrow i + 1$
    return 1
```

n is prime if the function returns 1. Let the length of the input be $\lceil\log_2 n\rceil = b$ bits. The number of iterations is $O(\sqrt{n}) = O(2^{\frac{b}{2}}) = 2^{O(b)}$. So the time complexity is exponential.

To get an idea about the problem, let us take n to be a 100 digit decimal number. The number of iterations in the worst case (when n is a prime) is $\sqrt{n} = 10^{50}$. If there are 10^9 iterations per second, the worst case time requirement is $10^{50-9}/(3.154 \times 10^7) \approx 10^{33}$ years. So the algorithm, though deterministic, is not suitable for testing a *large prime number*[1].

A modified trial division algorithm not only tests for primality, but also gives the prime decomposition or factoring. We may slightly improve the factoring algorithm by trial division. We divide n by 2 as many times as possible, so that $n = 2^k n_2$, where k is odd. Then n_2 will be divided with odd integers. But still in the worst case it is $O(\sqrt{n})$. After trying to divide with 3, if the quotient is $n_3 > 1$, the subsequent trials of division will be with integers that are congruent to 1 or 5 (mod 6). And this process will continue, until the quotient is 1.

One should look for characterisation of prime numbers that can be transformed into an efficient algorithm. A well known characterisation is the Wilson's theorem.

Theorem 1. (*Wilson*) p is a prime number if and only if $(p-1)! \equiv -1 \pmod{p}$.

This characterisation can be transformed to an efficient prime-testing algorithm if there is a fast computation method of $(n-1)! \bmod n$. Unfortunately no efficient procedure is known.

Another characterisation comes from a different area. The 10^{th} problem of the list of Hilbert problems[2] was as follows:

Is there any algorithm that, given a polynomial with integer coefficients, decides whether that polynomial has a root in integers?

This was proved to be *undecidable* by Yuri Matijasevič in 1970. There he demonstrated how to construct a polynomial of integer coefficients, of degree 10 with 26 variables, such that whenever integers are substituted for the variables,

[1] Large compared to what one come across every day. No given natural number can be called “large”, as there are infinitely many larger integers!

[2] Twenty-three mathematical problems were published by David Hilbert, a well known German mathematician, in 1900. According to Hilbert these are the problems that will keep the 20th century mathematicians busy. Many of these problems actually influenced the 20th century mathematics. A subset of theses problems (10 problems) were presented at the International Congress of Mathematicians in Paris.

the value of the polynomial is a prime number. And every prime number is a value of the polynomial.

But it is unknown, how to test efficiently whether a given positive integer is a value of the polynomial.

Now we consider an algorithm of very different flavor. It uses randomization. It has an *oracle* that can supply a "*random number*". In practical implementation we shall use "*pseudo random numbers*".

1.2 Lehmann's Randomized Algorithm

Following *randomized algorithm* for testing prime is based on Euler's criterion.

```
isprimeLR($n, t$)
    for $i \leftarrow 1$ to $t$ do
        $a \leftarrow rand\{1, 2, \cdots n-1\}$
        $r \leftarrow a^{\frac{n-1}{2}} \mod n$
        if $r \notin \{1, -1\}$ return 0
        else $b[i] \leftarrow r$
    if $\exists i,\ b[i] = -1$ then return 1
    else return 0
```

The algorithm in its i^{th} iteration chooses an integer a_i, *uniformly at random*, from the range $\{1, 2, \cdots, n-1\}$.

It evaluates $a_i^{\frac{n-1}{2}} \pmod n = r_i$ (say). If the value of r_i is anything other than 1 or $n-1$, the algorithm returns 0 indicating n as a composite number.

If the $r_i = 1$ or $n-1$, it is stored in the i^{th} element of the array $b[]$. The number of iterations t is also an input to the algorithm. If the outcome of all iterations are 1 or $n-1$, the computation comes to the next phase.

If all elements of the array are 1, the algorithm returns 0 i.e. reports n as a composite number. Otherwise, if there is at least one $n-1$, then it returns 1 and reports n as a prime number.

The algorithm is a finite sequence of random experiments. So the result is a random variable. We wish to know the probability of getting a "*wrong*" output. The output may be wrong in two ways. When n is prime and the algorithm reports it to be composite number, and when n is a composite number, the algorithm reports it as a prime.

When n is a prime number: half of the elements of $\{1, 2, \cdots, n-1\} = \mathbb{Z}_n^*$ should give the value of $a^{\frac{n-1}{2}} \bmod n$ as 1, and the other the other half should give the value as $n-1$ (Euler criterion).

So the loop will run for t iterations. And the probability of getting all t values of $a^{\frac{n-1}{2}}$ as 1 i.e. $r_1 = r_2 = \cdots = r_t = 1$ is $\frac{1}{2^t}$. So the probability of wrongly reporting n to be composite is also $\frac{1}{2^t}$.

When n is a composite: There are two possibilities. There may not be any $a \in \{1, 2, \cdots, n-1\}$ so that $a^{\frac{n-1}{2}} \bmod n = n-1$. Then either a value of $a^{\frac{n-1}{2}} \bmod n = d$, where $1 < d < n-1$, and the loop terminates without going through t iterations; or the loop is running for t iterations with all values of $a^{\frac{n-1}{2}} \bmod n = 1$. In both situations the algorithm correctly reports n to be composite.

The other possibility is that it will be reported as a *prime* after t iterations i.e. there is an $a \in \{1, 2, \cdots, n-1\}$ such that $a^{\frac{n-1}{2}} \bmod n = n-1$. In that case there is a theorem which claims that for more than half of $a \in \{1, 2, \cdots, n-1\}$, $a^{\frac{n-1}{2}} \bmod n$ is neither 1 nor $n-1$.

In this case the probability of the loop running for t iterations is not more than 2^{-t}. So, the probability that n will be reported as a prime (wrongly), at the end of t iterations, cannot be more than 2^{-t}.

In both the cases, the probability of wrong output is bounded by 2^{-t}. This can be made low by increasing the number of iterations.

The main cost is the computation of $a^{\frac{n-1}{2}} \bmod n$. The size of any intermediate data is bounded by n^2 (two elements of $\{1, 2, \cdots, n-1\}$ are multiplied). The exponentiation can be done by *repeated squaring* that requires $O(\log_2 n)$ iterations.

Following is the algorithm for computing of $a^e \bmod n$ where $a \in \mathbb{Z}_n$ and e is a positive integer. Let the binary representation of $e = (e_{k-1}e_{k-2} \cdots e_1e_0)$.

```
modExpN($a, e, n$)
    $exp \leftarrow 1$
    $s \leftarrow$ a mod $n$
    while $e \geq 1$
        if ($e$ mod 2) $= 1$ then $exp \leftarrow (exp \times s)$ mod $n$
        $s \leftarrow s^2$ mod $n$
        $e \leftarrow e \div 2$
    return $exp$
```

$k = \lceil \log_2 e \rceil$, so the loop is executed k times with k *squaring* and $\leq k$ multiplications over $\mathbb{Z}_n$. So the running time is $O(\log e (\log n)^2)$. If $1 < e < n$, then it is $O(\log n)^3)$.

1.3 Fermat Test

Fermat's little theorem states that: if p is a prime and $p \not| a$, then $a^{p-1} \equiv 1 \pmod{p}$.
In other words, if $gcd(a, p) = 1$, and $a^{p-1} \not\equiv 1 \pmod{p}$, then p is composite.
We may choose a from $\mathbb{Z}_n$.
Definition 1: An integer a, $1 \leq a < n$, is called a *Fermat-witness* or *F-witness* for n as a composite number, if a^{n-1} mod $n \neq 1$.

Example 1. 2^{1246} mod $1247 = 173$, so 2 is an F-witness for 1247 (as composite). On the other hand 2^{340} mod $341 = 1$, but $341 = 11 \times 31$, is a composite. So 2 is not an F-witness for 341. But 3^{340} mod $341 = 56$, so 3 is an F-witness for 341.
Definition 2: An integer $a \in (1, n)$, is called a *Fermat-liar* or *F-liar* for an odd composite number n, if a^{n-1} mod $n = 1$.

2 is a F-liar for 341. The converse of Fermat's little theorem is not true in general. But it is almost true as the number of F-liar are small for most composite numbers.
Proposition 2. Let n be an integer and $n \geq 2$.

1. If $1 \leq a < n$ is such that $a^r \equiv 1 \pmod{n}$ for some integer $r \geq 1$, then $a \in \mathbb{Z}_n^*$.
2. If $a^{n-1} \equiv 1 \pmod{n}$ for all a, $1 \leq a < n$, then n is a prime number.
3. If n is an odd composite number, then every element of $\mathbb{Z}_n \setminus (\mathbb{Z}_n^* \cup \{0\})$ is an *F-witness*.

Proof:

1. If $r = 1$, then $a \equiv 1 \pmod{n}$, implies that $n | a - 1$. But $a < n$, so $a = 1 \in \mathbb{Z}_n^*$.
   Otherwise, $a \cdot a^{r-1} \equiv 1 \pmod{n}$. Let us call $a^{r-1} = b$. So b is the multiplicative inverse of a and both of them are in $\mathbb{Z}_n^*$. Note that $ax \equiv 1 (\text{mod } n)$ has a solution implies that $gcd(a, n) = 1$ i.e. $a \in \mathbb{Z}_n^*$.

   In other words, $a \cdot a^{r-1} + nk = 1$. So by the Bezout's identity, $gcd(a, n) = 1$ and $a \in \mathbb{Z}_n^*$.
2. We have $1^{n-1} \equiv 2^{n-1} \equiv \cdots \equiv (n-1)^{n-1} \equiv 1 \pmod{n}$.
   From the first part of the proposition we have $1, 2, \cdots, n-1$ are elements of $\mathbb{Z}_n^*$. This is equivalent to say that n is a prime.
3. If $a \in \mathbb{Z}_n \setminus \mathbb{Z}_n^*$, then $gcd(a, n) = d > 1$. If $a^{n-1} \equiv 1 \pmod{n}$, then $a \cdot a^{n-2} + kn = 1$, which is impossible as $d > 1$ and d is the smallest positive integer that can be expressed as linear combination (diophantine) of a and n. So $a^{n-1} \not\equiv 1 \pmod{n}$ and is an *F-witness*.

QED.

An F-witnesses from $\mathbb{Z}_n \setminus (\mathbb{Z}_n^* \cup \{0\})$ may be used to prove that n is composite. But the number elements in $\mathbb{Z}_n \setminus (\mathbb{Z}_n^* \cup \{0\})$ is $n - 1 - \phi(n)$ is small[3] compared to $\{1, \cdots, n-1\}$. So the probability of selecting them at random is low.

[3] Consider $n = pq$, where p, q are almost equal size primes. The size of $\mathbb{Z}_n \setminus \mathbb{Z}_n^*$ is $n - 1 - \phi(n) = (pq - 1) - (p - 1)(q - 1) = p + q - 2$.

Example 2. We consider $35 = 5 \times 7$, an odd composite number with two prime factors. There are $35 - 1 - \phi(35) = 34 - (5-1)(7-1) = 11$ elements in $\mathbb{Z}_{35} \setminus \mathbb{Z}^*_{35}$. Out of which $11 - 1 = 10$ (excluding 0) are F-witnesses.

But in most of cases, there are elements of $\mathbb{Z}^*_n$ that are also F-witnesses.

Example 3. Consider the previous example of $5 \times 7 = 35$

$\mathbb{Z}_{35} \setminus (\mathbb{Z}^*_{35} \cup \{0\})$	F-witness in $\mathbb{Z}^*_{35}$	F-liar in $\mathbb{Z}^*_{35}$
5, 7, 10, 14, 15, 20, 21, 25, 28, 30	2, 3, 4, 8, 9, 11, 12, 13, 16, 17, 18, 19, 22, 23, 24, 26, 27, 31, 32, 33,	1, 6, 29, 34

So it seems that for an odd n an a chosen *uniformly at random* from $\{2, \cdots, n-2\}$ will have larger than $\frac{1}{2}$ probability of being an *F-witness.*
1 and $n - 1$ are always *F-liars* for an odd n as $n - 1 \equiv -1 \pmod{n}$ and $(-1)^{n-1} = 1$. So we choose *a uniformly at random* from $\{2, 3, \cdots, n-2\}$. In this example the the probability of getting an *F-liar* is $\frac{2}{32} = \frac{1}{16}$. This is a good news.

The first approximation of primality test based on the above fact is known as Fermat test. It depends on *F-witnesses.*

```
isprimeFT1($n$) // n is odd $\geq 3$
    $a \leftarrow$ rand$\{2, \cdots n-2\}$
    if ($a^{n-1}$ mod $n$) $\neq 1$ return 0 // composite
    else return 1 // prime
```

Proposition 3. Let $n \geq 3$ be an odd composite number. If there is an F-witness $a \in \mathbb{Z}^*_n$, then `isprimeFT1()` returns 0, a proof of composite number, with a probability more than $\frac{1}{2}$.
Proof: We define the set of F-liars for odd composite n as

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

It is clear from the *proposition* (1.3.1) that $L^F_n \subseteq \mathbb{Z}^*_n$. We prove that it is a subgroup of $\mathbb{Z}^*_n$. It is enough to show that L^F_n is closed under $\times_n$ as $\mathbb{Z}^*_n$ is finite.
Closure: if $a, b \in L^F_n$, then $a^{n-1} \equiv 1 \pmod{n}$ and $b^{n-1} \equiv 1 \pmod{n}$, so $(ab)^{n-1} = a^{n-1} \cdot b^{n-1} \equiv 1 \pmod{n}$. So $ab \in L^F_n$.

If there is one *F-witness* in $\mathbb{Z}^*_n$, the set of *F-liar*, L^F_n, cannot be equal to $\mathbb{Z}^*_n$ and is a *proper* subgroup of $\mathbb{Z}^*_n$. By Lagrange's theorem the order of L^F_n divides the order of $\mathbb{Z}^*_n = \phi(n) < n - 1$, as n is a composite number. So the size or order of L^F_n, $|L^F_n| < \frac{n-1}{2}$.

Thus the probability that a is chosen uniformly at random from $\{2, \cdots, n-2\}$ will be an *F-liar* is at most

$$\frac{|L^F_n \setminus \{1, n-1\}|}{|\mathbb{Z}_n \setminus \{0, 1, n-1\}|} < \frac{\frac{n-1}{2} - 2}{n-3} = \frac{n-5}{2(n-3)} < \frac{1}{2}.$$

QED.

Under the given assumption (that there is at least one *F-witness* in $\mathbb{Z}^*_n$), the probability of getting *F-liars* after k trails or *random experiment* of *primality test* is less than 2^{-k}. Following is the k-iteration *Fermat test.*

```
isprimeFT2($n, k$) // n is odd $\geq 3$ $k \geq 1$
    for $i \leftarrow 1$ to $k$
        $a \leftarrow rand\{2, \cdots n-2\}$
        if a$^{n-1}$mod $n \neq 1$ return 0 // composite
    else return 1 // prime
```

The algorithm returns 0 (n is composite) if it has found an F-witness. If n is a prime, no such witness is possible to get so the algorithm returns 1.

1.3.1 Carmichael Numbers

But the story does not end here. What if there is some odd composite n so that there is no F-witness in $\mathbb{Z}_n^*$. Interestingly there are such species, and in fact there are infinitely many of them (proved in 1994).

Definition 3: An odd composite number n is called a *Carmichael number* if for all $a \in \mathbb{Z}_n^*$, $a^{n-1} \equiv 1 \pmod{n}$ i.e. all elements of $\mathbb{Z}_n^*$ are F-liars.

In other words, an odd composite number n is called a *Carmichael number* if $a^n \equiv a \pmod{n}$, for all integer a.

The smallest Carmichael number is $561 = 3 \cdot 11 \cdot 17$. All elements of $\mathbb{Z}_{561}^*$ ($\phi(561) = 561 \cdot \frac{2}{3} \cdot \frac{10}{11} \cdot \frac{16}{17} = 320$) are *F-liars* and the probability of getting wrong answer is $\frac{320-2}{561-3} = 0.5699 > 0.5$.

In general, if a Carmichael number is passed through the Fermat Test, the probability that it will be wrongly identified as a prime is

$$\frac{|\mathbb{Z}_n^* \setminus \{1, n-1\}|}{|\{2, 3, \cdots, n-2\}|} = \frac{\phi(n) - 2}{n - 3} > \frac{\phi(n)}{n} = \prod_{\substack{\text{prime } p \\ p|n}} \left(1 - \frac{1}{p}\right).$$

This bound may be close to 1 for Carmichael numbers with large prime factors. As an example for a Carmichael number $651693055693681 = 72931 \times 87517 \times 102103$, the probability is greater than 0.999965.

Let $C(x)$ be the number of Carmichael numbers below x. It was proved that $C(x) > x^{\frac{2}{7}}$. The bound was improved in 2005 to $x^{0.332}$. Following table is from Wikipedia showing the initial distribution.

n	3	4	5	6	7	$\cdots$	20	$\cdots$
$C(10^n)$	1	7	16	43	105	$\cdots$	8220777	$\cdots$

So the conclusion is that we have to go beyond Fermat test if we want to test primality with reliability. Following are a few properties of Carmichael numbers.

Proposition 4. A Carmichael number cannot be of the form p^k, where p is an odd prime and $k > 1$.

Proof: Our basic strategy is to find an appropriate element of $\mathbb{Z}_{p^k}^*$ that is an F-witness.

Let $n = p^k$ be a Carmichael number, where p is an odd prime. We choose $a = p + 1$. The $gcd(a, p) = 1$ implies that $gcd(n, a) = 1$. So $a \in \mathbb{Z}_n^*$.

We assume that a is an F-liar, so $a^{n-1} \equiv 1 \pmod{n}$, and that implies $a^{n-1} \equiv 1 \pmod{p^2}$ i.e.

$$a^{n-1} \equiv (1+p)^{n-1} \equiv 1 + (n-1)p + \sum_{2 \le i \le n-1} \binom{n-1}{i} p^i \equiv 1 + (n-1)p \pmod{p^2}.$$

So $(n-1)p \equiv a^{n-1} - 1 \equiv 0 \pmod{p^2}$ i.e. $p|(n-1)$, but it is impossible as $n = p^k$. So a is an *F-witness* of n in $\mathbb{Z}_n^*$. Hence n cannot be a Carmichael number. QED.

Proposition 5. If n is a Carmichael number then n is a product of at least three distinct primes, and none of them have power more than 1.

Proof: We have already proved that n cannot be of the form p^k, where p is an odd prime and $k > 1$.

There are two other cases to consider:

1. $n = p^k \times m$, where $m > 1$, p is a prime, $k \ge 2$, and $gcd(p, m) = 1$.
2. $n = p \times q$, where p and q are primes.

We shall show that in each of these cases there is an *F-witness* in $\mathbb{Z}_n^*$.

Case 1: $n = p^k \times m$: clearly p^2 and m are relatively prime. According to the Chinese remainder theorem we can get an a such that $1 \le a < p^2 \times m \le n$ and satisfies the congruence

$$\begin{aligned} a &\equiv p + 1 \pmod{p^2}, \text{ and,} \\ a &\equiv 1 \pmod{m}. \end{aligned}$$

Clearly $gcd(a, p^k) = 1$ as $p^2|a - (p+1)$ implies $p|a - (p+1)$ implies $p|a - 1$, so $p \not| a$. Also $gcd(a, m) = 1$.
So the $gcd(a, n) = 1$ i.e. $a \in \mathbb{Z}_n^*$.
Our next claim is that a is an F-witness. If it is not the case, we have $a^{n-1} \equiv 1 \pmod{n}$ implies $a^{n-1} \equiv 1 \pmod{p^2}$ as $p^2|n$.
Again,

$$1 \equiv a^{n-1} \equiv (1+p)^{n-1} \equiv 1+(n-1)p + \sum_{2 \leq i \leq n-1} \binom{n-1}{i} p^i \equiv 1+(n-1)p \ (\bmod\ p^2).$$

So we have $(n-1)p = (p^k \cdot m - 1)p \equiv 0 \pmod{p^2}$ i.e. $p^k \cdot m - 1 \equiv 0 \pmod{p}$. But that is impossible. So a is an F-witness of n.
Case 2: $n = p \times q$. Without any loss of generality we assume that $p > q$.

It is known that $\mathbb{Z}_p^*$ is a cyclic group and it has a generator say $g < p$. Again using the Chinese remainder theorem there is an element a, $1 \leq a < n$ satisfying

$$\begin{aligned} a &\equiv g \pmod{p}, \\ a &\equiv 1 \pmod{q}. \end{aligned}$$

It is clear that $gcd(a, p) = gcd(a, g) = 1$ and $gcd(a, q) = 1$. So $a \in \mathbb{Z}_n^*$.
We claim that a is an *F-witness*. Otherwise (if a is an *F-liar*), $a^{n-1} \equiv 1 \pmod{n}$, implies that $a^{n-1} \equiv 1 \pmod{p}$. But then $a \equiv g (\bmod\ p)$, so $g^{n-1} \equiv 1 \pmod{p}$.

As g is the generator of $\mathbb{Z}_p^*$, $p-1$ is the order of g, so $(p-1)|(n-1)$ i.e. $(p-1)|(pq-1)$, implies $(p-1)|[q(p-1)+q-1]$, implies $(p-1)|(q-1)$, implies $p \leq q$ - a contradiction. QED.

1.3.2 Square Roots of 1

Definition 4: Let n be a positive integer and a is an integer such that $1 \leq a < n$. a is called *square root of 1 modulo* n if $a^2 \equiv 1 \pmod{n}$.

For every n we have $1^2 = 1$ and $(n-1)^2 \equiv (-1)^2 = 1 \pmod{n}$. These are known as *trivial* square roots of 1 modulo n.
We have already proved that $x^2 \equiv 1 (\bmod\ p^e)$, where p is an odd prime and e is a positive integer has only these two solutions, $x = 1$ or $x = p^e - 1 \equiv -1 \pmod{p^e}$.
We also have seen that for an odd positive integer n whose prime factorisation is

$$n = p_1^{e_1} \cdot \cdots \cdot p_k^{e_k},$$

where $k \geq 2$, has 2^k solutions of $a^2 \equiv 1 \pmod{n}$. The solutions are pre-images of $(\{1, p_1^{e_1} - 1\}, \cdots, \{1, p_k^{e_k} - 1\})$ in the Chinese Remainder map

$$f : \mathbb{Z}_n^* \to \mathbb{Z}_{p_1^{e_1}}^* \times \cdots \times \mathbb{Z}_{p_k^{e_k}}^*.$$

So for an odd composite integer n there are *non-trivial* square-roots of 1.

Example 4. If $n = 977$ a prime, it has only two square-roots of 1 modulo 977. They are $1, 976$.
If $n = 125$, then also 1 and 124 are the only two square-roots of 1 modulo 125. But if $n = 7875 = 3^2 \times 5^3 \times 7$, there are 8 square-roots of 1 modulo 7875. These are $f^{-1}(1,1,1) = 1$, $f^{-1}(8,1,6) = 251$, $f^{-1}(1,124,6) = 874$, $f^{-1}(1,1,6) = 1126$, $f^{-1}(8,124,1) = 6749$, $f^{-1}(8,1,1) = 7001$, $f^{-1}(1,124,1) = 7624$ and $f^{-1}(8,124,6) = 7874$.

It is not difficult (computationally) to detect a *perfect power*. Let n be an odd positive integer which is not a *perfect power*. If it has *non-trivial* square root of 1 modulo n, then it is a composite number. But this property alone cannot be used in a randomized algorithm unless n has many prime factors.

We introduce this test for non-trivial square root of 1 in the process of Fermat test. In Fermat's test we compute $a^{n-1} \bmod\ n$, where n is an odd positive integer. So $n-1$ is even, and can be written as, $n - 1 = u \cdot 2^k$, where u is an odd integer. We can write

$$a^{n-1} \equiv (a^u \bmod n)^{2^k} \pmod{n}.$$

So we have $k+1$ steps of computation where,

$$b_i = \begin{cases} a^u \bmod n & \text{if } i = 0, \\ (b_{i-1})^2 \bmod n & \text{if } i = 1, 2, \cdots, k. \end{cases}$$

Example 5. Let $n = 3601 = 13 \cdot 277$, we have $n - 1 = 3600 = 2^4 \cdot 225$. We have the following table for a few values of a, $2 \leq a \leq 3599$.

a	$b_0 = a^{225} \bmod n$	$b_1 = b_0^2 \bmod n$	$b_2 = b_1^2 \bmod n$	$b_3 = b_2^2 \bmod n$	$b_4 = b_3^2 \bmod n$
9	2380	27	729	2094	2419
13	2015	1898	1404	1469	962
35	2276	1938	1	1	1
60	60	3600	1	1	1
278	278	1663	1	1	1
555	1	1	1	1	1

It is clear that 9 and 13 $\in \mathbb{Z}_{3601} \setminus \mathbb{Z}^*_{3601}$ are *F-witnesses* showing 3601 as composite. Both 35, 60, 278 and 555 are *F-liars*. But in the rows of 35 and 278 we observe that 1938 and 1663 are non-trivial square roots of 1 modulo 3601. So both of them indicate that 3601 is a composite number. No such conclusion can be drawn from other two. The conclusion drawn from the row of 35 and the row of 278 are beyond the *Fermat test*.

The prime factorisation of $3601 = 13 \times 277$. So there are $2 \times 2 = 4$ square roots of 1. They are $(1,1) \mapsto 1$, $(1,276) \mapsto 1938$, $(12,1) \mapsto 1663$ and $(12,276) \mapsto 3600$. Out of these four 1663 and 1938 are non-trivial.

For a given a the sequence of $b_0, b_1, \cdots, b_k$ may have four possible forms.

1. $b_0 = 1$: The complete sequence will have 1's. The value of a may be an F-liar, and it does not show any nontrivial square root of 1. No definite conclusion can be drawn about n.

2. $b_0 \neq 1$ but $b_i = n - 1$ for some $i < k$: The sequence $b_{i+1}, \cdots, b_k$ is of 1's. Again a may be an F-liar, and there is no nontrivial square root of 1 in the sequence. No definite conclusion can be drawn about n.

3. $b_0, \cdots, b_k \in \{2, \cdots, n-2\}$ - a is an F-witness and n is certainly composite.

4. $b_0 \neq 1$ and none of $b_0, b_1, \cdots, b_{k-1}$ are $n - 1$, but $b_k = 1$: So there is a smallest i, $1 \leq i \leq k$ so that $b_i = 1$. This implies that $b_{i-1}^2 \equiv 1 \pmod{n}$, i.e. b_{i-1} is a non-trivial square root of 1. So n is composite.

Last two cases detect composite. We can summaries these two cases as follows.
If $a^u \bmod\ n \neq 1$ and $a^{u \cdot 2^i} \bmod\ n \neq n - 1$ for $i = 0, 1, \cdots, k-1$, then n is a composite number. This decision actually does not depend on whether b_k is 1 or not.
Definition 5: Let $n \geq 3$ be an odd integer so that $n - 1 = u \cdot 2^k$, where u is an odd integer and $k \geq 1$. An integer a, $1 \leq a < n$, is called an *Artjuhov-witness* or *A-witness* of n to be composite, if $a^u \not\equiv 1 \pmod{n}$ and $a^{u \cdot 2^i} \not\equiv -1 \pmod{n}$ for all i, $0 \leq i < k$. If n is composite and a is not an *A-witness* for n, then a is called an *A-liar*.

We define the set of A-liars for an odd composite n where $n - 1 = u \times 2^k$.

$$L_n^A = \{a \in \mathbb{Z}_n : b_0 = a^u \bmod\ n = 1 \text{ or } b_i \bmod\ n = n - 1,\ 0 \leq i < k\}.$$

It is important to note that an integer a, $1 \leq a < n$, satisfying the cases 1 or 2 or 4 is an element of L_n^F. Whereas an a satisfying cases 1 or 2 is an element of L_n^A. So $L_n^A \subseteq L_n^F$. So for a given odd integer n, the set of *A-witness* may be a larger set.
Proposition 6. If a is an *A-witness* for n, then n is composite.

The strengthening of Fermat test using A-witness is called *Miller-Rabin test*[4]. Following is the randomised *Miller-Rabin* algorithm.

[4] M. Artjuhov [AM] suggested the sequence of computation of b_i. Afterward G. M. Miller [MGM] used this criteria and proposed a deterministic polynomial time algorithm under the assumption that the Extended Riemann Hypothesis (ERH) is true. He proved that the small-

```
isprimeRM($n$) // n is odd $\geq 3$
1   Find $u$, $k$ so that $n-1 = u \cdot 2^k$
2   $a \leftarrow rand\{2, \cdots n-2\}$
3   $b \leftarrow a^u \bmod n$
4   if $b \in \{1, n-1\}$ return 1
5   for $i \leftarrow 1$ to $k-1$
6       $b \leftarrow b^2 \bmod n$
7       if $b = n-1$ return 1
8       if $b = 1$ return 0 // certainly composite
9   return 0 // certainly composite
```

Extraction of u and k, where $n = u \cdot 2^k$ takes at most $\log n$ division by 2. This is equivalent to right-shift operations and takes $O((\log n)^2)$ bit operations. Computation of $a^u \bmod n$ takes $O((\log n)^3)$ bit operations. The loop is executed $k-1$ times which is $O(\log n)$. Multiplication modulo n is the most costly operation per iteration. It takes $O((\log n)^2)$ bit operations. So there are $O(\log n)$ arithmetic operations and $O((\log n)^3)$ bit operations.

Lemma 7. If the Miller-Rabin test returns 0, then n is composite.

Proof: Let a be chosen in the algorithm and the output is 0. We claim that a is an A-witness and n is a composite number.

Value 0 may be returned in two different ways:

Case I: We get 0 due to return from line 8 in the i^{th} iteration i.e. $b_i = 1$. It is clear that $b_0, b_1, \cdots, b_{i-1} \notin \{1, n-1\}$. So b_{i-1} must be a non-trivial square root of 1. So n is composite and a is an A-witness.

Case II: We get a 0 due to return from line 9 at the end of $k-1$ iterations. So $b_0, b_1, \cdots, b_{k-1} \notin \{1, n-1\}$. So by the definition a is an A-witness that n is a composite. QED.

It remains to show that the algorithm `isprimeMR()` gives an erroneous output 1 for an odd composite n with a *bounded probability*, so that we can iterate it to any desired low value.

We wish to demonstrate that the size of A-liars is less than half of the size of $\{2, 3, \cdots, n-2\}$. But unlike the set L_n^F, the set L_n^A is not a subgroup of $\mathbb{Z}_n^*$. In fact it is not closed under the modulo n multiplication. Product of two elements of L_n^A may not belong to L_n^A.

Example 6. Consider the previous example $n = 3601 = 13 \cdot 277$.

a	$b_0 = a^{225} \bmod n$	$b_1 = b_0^2 \bmod n$	$b_2 = b_1^2 \bmod n$	$b_3 = b_2^2 \bmod n$	$b_4 = b_3^2 \bmod n$
60	60	3600	1	1	1
242	1048	3600	1	1	1
116	1663	1	1	1	1

Note that both 60 and 242 are *A-liars*, but their product $(60 \times 242) \bmod 3601 = 116$ is an *A-witness*.

We shall try to identify a proper subgroup of $\mathbb{Z}_n^*$, so that L_n^A is a subset of it. If n is not a Carmichael number then L_n^F is a proper subgroup of $\mathbb{Z}_n^*$. It is also a superset of L_n^A. So our earlier argument of getting a bounded probability for *false 1* ($< 1/2$) works. Unfortunately if n is a Carmichael number, $L_n^F = \mathbb{Z}_n^*$, and the argument does not stand. So we search for a different subgroup.

We assume that n is a Carmichael number. We try to find B_n^A, a proper subgroup of $\mathbb{Z}_n^*$, so that $L_n^A \subseteq B_n^A$ and our previous argument about size of L_n^A works (order is half of $\mathbb{Z}_n^*$).

Definition 6: Let n be a Carmichael number and let $i_0 \geq 0$ be the largest integer so that there exists an *A-liar* a_0 such that $a_0^{u \cdot 2^{i_0}} \equiv -1(\bmod\ n)$ $(1 \leq a_0 < n$ and $n-1 = u \cdot 2^k)$. Such an i_0 exists as if we take $a = n-1$, then $(n-1)^{u \cdot 2^0} \equiv (-1)^u \equiv -1 \ (\bmod\ n)$.

As n is a Carmichael number and a_0 is an *A-liar*, we have $1 \equiv a_0^{n-1} \equiv a_0^{u \cdot 2^k} (\bmod\ n)$. So the value of i_0 must be less than k, i.e. $0 \leq i_0 < k$. We define

est *A-witness* of a composite number n is of size $O(\ln n)^2)$ provided ERH is true. Subsequently E. Bach [BE] gave an explicit bound of $2(\ln n)^2$ of the smallest *A-witness* under the same assumption. The algorithm is simple, it tries with $a = 2, \cdots, \lfloor 2(\ln n)^2 \rfloor$ as *A-witness*. If all of them fails, then n is prime. Subsequently M. O. Rabin [RMO] proposed the randomized algorithm on *A-witness*.

the following set using this i_0.

$$B_n^A = \{a : 1 \leq a < n \text{ and } a^{u \cdot 2^{i_0}} \equiv \pm 1 (\bmod\ n)\}.$$

The following proposition will establish that $L_n^A \subseteq B_n^A$ and B_n^A is a proper subgroup of $\mathbb{Z}_n^*$. So its size is less than or equal to half of $\mathbb{Z}_n^*$[5].

Proposition 8.

1. $L_n^A \subseteq B_n^A$,
2. B_n^A is a subgroup of $\mathbb{Z}_n^*$,
3. $\mathbb{Z}_n^* \setminus B_n^A \neq \emptyset$ i.e. B_n^A is a proper subgroup.

Proof:

1. Let $a \in L_n^A$. There are two possible cases from the definition of L_n^A:
   Casr I: if $a^u \equiv 1 (\bmod\ n)$, then $a^{u \cdot 2^{i_0}} \equiv 1 (\bmod\ n)$, so $a \in B_n^A$.
   Case II: if $a^{u \cdot 2^i} \equiv -1 (\bmod\ n)$, then there are two possible cases. If $i = i_0$, then $a \in B_n^A$. Otherwise, $i < i_0$, so $a^{u \cdot 2^{i_0}} \equiv (a^{u \cdot 2^i})^{2^{i_0 - i}} \equiv (-1)^{2^{i_0 - i}} \equiv 1 (\bmod\ n)$. So $a \in B_n^A$.

2. We check that B_n^A is closed under the group operation. Let $a, b \in B_n^A$, then $a^{u \cdot 2^{i_0}} \equiv \pm 1 (\bmod\ n)$ and $b^{u \cdot 2^{i_0}} \equiv \pm 1 (\bmod\ n)$.

   So $(ab)^{u \cdot 2^{i_0}} \equiv (a^{u \cdot 2^{i_0}})(b^{u \cdot 2^{i_0}}) \equiv (\pm 1)(\pm 1) \equiv \pm 1 (\bmod\ n)$.

   We conclude that B_n^A is a subgroup of $\mathbb{Z}_n^*$.

3. We have already proved that a Carmichael number cannot be a power of one prime. So $n = n_1 \times n_2$, where n_1 and n_2 are relatively prime odd numbers.

   Let a_0 be an A-liar such that $a_0^{u \cdot 2^{i_0}} \equiv -1 (\bmod\ n)$ (there is always one). We have an a_1, $1 \leq a_1 < n_1$ and $a_1 \equiv a_0 \pmod{n_1}$. By the Chinese Remainder Theorem there is a unique $a \in \mathbb{Z}_n$ such that

   $$\begin{aligned} a &\equiv a_1 \pmod{n_1}, \\ a &\equiv 1 \pmod{n_2}. \end{aligned}$$

   We claim that $a \in \mathbb{Z}_n^* \setminus B_n^A$.
   We have $a_1 \equiv a_0 \pmod{n_1}$ and $a \equiv a_1 \pmod{n_1}$, implies that $a \equiv a_0 \pmod{n_1}$. So we have $a^{u \cdot 2^{i_0}} \equiv a_0^{u \cdot 2^{i_0}} \equiv -1 \pmod{n_1}$.

   Similarly we calculate $a^{u \cdot 2^{i_0}} \equiv 1^{u \cdot 2^{i_0}} \equiv 1 \pmod{n_2}$.

   So $a^{u \cdot 2^{i_0}} \bmod n$ is neither equal to 1 nor equal to -1. So $a \notin B_n^A$.

   But then $a^{u \cdot 2^{i_0 + 1}} \equiv 1 \pmod{n_1}$ and also $a^{u \cdot 2^{i_0+1}} \equiv 1 \pmod{n_2}$. So we have $a^{u \cdot 2^{i_0+1}} \equiv 1 \pmod{n_1 n_2}$ as the $gcd(n_1, n_2) = 1$. So $a \in \mathbb{Z}_n^*$.

QED.

We have proved that B_n^A is a proper subgroup $\mathbb{Z}_n^*$ and $L_n^A \subseteq B_n^A$. The order of $B_n^A \leq \frac{\phi(n)}{2}$. A Carmichael number n has at least three prime factors. So the value of $\phi(n)$ is much less than $n-2$. And the probability of a randomly chosen a from $\{2, \cdots, n-2\}$ to be in L_n^A is less than $\frac{(n-2)/2-2}{n-3} < \frac{1}{2}$.

A better analysis can prove that the size of *A-liar* set $|L_n^A| \leq |\mathbb{Z}_n^*|/4$. If we iterate the algorithm for k times, the probability that the *Rabin-Miller* algorithm will declare a composite n to be prime is

$$\gamma(n, k) \leq \frac{n-1}{\pi(n)} 4^{-k} = O(l 4^{-k}),$$

where $\pi(n)$ be the set of primes that are $\leq n$ is estimated to be $\geq c(n-1)/\log n$. So l is the length of n.

Iterated version of *Rabin-Miller* algorithm.

[5] This also answers about the condition of the Lehmann's algorithm which says that, if there is some $a \in \{1, \cdots, n-1\}$ so that $a^{\frac{n-1}{2}} \equiv a^{u \cdot 2^{k-1}} \equiv -1 (\bmod\ n)$, then more than half of the elements b of $\{1, \cdots, n-1\}$ satisfy $b^{\frac{n-1}{2}} \equiv b^{u \cdot 2^{k-1}} \not\equiv \pm 1 (\bmod\ n)$.

```
isprimeRM($n, l$) // n is odd $\geq 3$, $l \geq 1$
Find $u$, $k$ so that $n - 1 = u \cdot 2^k$
for $j \leftarrow 1$ to $l$
    $a \leftarrow rand\{2, \cdots n - 2\}$
    $b \leftarrow a^u \bmod n$
    if $b \in \{1, n - 1\}$ continue
    for $i \leftarrow 1$ to $k - 1$
        $b \leftarrow b^2 \bmod n$
        if $b = n - 1$ continue
        if $b = 1$ return 0 // certainly composite
    return 0 // certainly composite
return 1
```

1.3.3 Generating Random Prime

It is important to generate random prime for different cryptographic applications. Miller-Rabin algorithm is used to generate that. Following is a scheme to generate a random prime in the range $2, \cdots, m$.

```
randPrime($m, l$) // random prime $2 \cdots m$
do
    do
        $n \leftarrow rand\{2, \cdots m\}$
        if $n \equiv 0 (\text{mod } 2)$ continue
    while $\neg$isprimeRM($n, 1$)
while $\neg$isprimeRM($n, l$)
return $n$
```

The first test will almost often detect a composite n. The second test is to increase confidence.

In some application it is necessary to generate a random prime of certain length l e.g. a random prime of 1024 bits. This essentially means that the random prime $p \in [2^{l-1}, 2^l)$.

According to *Bertrand's Postulate*, for ever $n \geq 1$, there is a prime number p with $n < p \leq 2n$. A stronger result is that $\pi(2n) - \pi(n) > \frac{n}{3 \log 2n}$, where $\pi(x)$ is the number of primes less than or equal to x. It implies that $\pi(2^l) - \pi(2^{l-1}) \geq \frac{c2^{l-1}}{l}$ for $l \geq 2$. So there are large number of primes in the range. We generate random l-bits and use Millar-Rabin to test whether it is prime.

Example 7. $\pi(1024) - \pi(512) > \frac{512}{3 \log 1024} \approx 56$. There are actually $172 - 97 = 75$ primes in that range.

References

[AB] *Computational Number Theory* by Abhijit Das, (will be published from CRC Press).

[AM] *Certain criteria for the primality of numbers connected with the Little Fermat Theorem (in Russian)* by Artjuhov M, Acta. Arith., 12 (1966/67) 355-364.

[BE] *Explicit bounds for primality testing and related problems* by Bach E, Inform. and Comput. 90 (1990) 355-380.

[CLR] *Introduction to Algorithms* by Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein, 2nd ed., Pub. Pub. PHI, 2001, ISBN 81-203-2141-3.

[LKP] `http://primes.utm.edu/largest.html#largest`

[MD] *Primality Testing in Polynomial Time From Randomized Algorithms to "PRIMES is in P"* by Martin Dietzfelbinger, LNCS 3000, Pub. Springer, 2004, ISBN 3-540-40344-2.

[MGM] *Riemann's hypothesis and the tests for primality* by Miller G M, J. Comput. Syst. Sci. 13 (1976) 300-317.

[RMO] *Probabilistic algorithm for testing primality* by Rabin M O, J. Number Theory 12 (1980) 128-138.

[VS] *A Computational Introduction to Number Theory and Algebra* by Victor Shoup, 2nd ed., Pub. Cambridge University Press, 2009, ISBN 978-0-521-51644-0.