

Primality Tests

Andreas Klappenecker

Texas A&M University

© 2018 by Andreas Klappenecker. All rights reserved.

Suppose that Bob chooses the large number such as

$$n = 456\,989\,977\,669$$

How can Bob check whether n is prime?

Actually, this is a “small number”. We are usually interested in testing primality of numbers with hundreds of digits, but those do not look too nice on a slide.

- Use the Agrawal-Kayal-Saxena primality test.
- It is a deterministic $\tilde{O}(\log(n)^{12})$ time algorithm.
- Space requirements make the test impractical for large n .

Unlike integer factorization into primes, we know a poly-time algorithm for primality testing, but it is not too useful in practice.

Goal

We will now develop some randomized algorithms for primality testing.

Fermat

We need the following simple result from number theory.

Fermat's Little Theorem

Let p be a prime. Then

$$a^p \equiv a \pmod{p}$$

for all integers a .

Case 1. Suppose that p divides a . Then $a \equiv 0 \equiv a^p \pmod{p}$.

Case 2. Suppose that p does not divide a .

Then $a^p \equiv a \pmod{p}$ is equivalent to $a^{p-1} \equiv 1 \pmod{p}$.

Consider the $p - 1$ numbers

$$a, 2a, 3a, \dots, (p - 1)a.$$

We claim that they are all **different** mod p . Indeed, if we would have $ja \equiv ka \pmod{p}$, then $(j - k)a \equiv 0 \pmod{p}$. Since $a \not\equiv 0 \pmod{p}$, we must have $(j - k) \equiv 0 \pmod{p}$. So $j \equiv k \pmod{p}$. However, this implies $j = k$, since $1 \leq j, k < p$.

Proof of Fermat's Little Theorem (2/2)

Since $a, 2a, 3a, \dots, (p-1)a$ are $p-1$ different nonzero numbers mod p , we have

$$\{a, 2a, 3a, \dots, (p-1)a\} = \{1, 2, 3, \dots, p-1\} \pmod{p}$$

Multiplying these numbers together, we can conclude that

$$a^{p-1}(p-1)! = a \cdot 2a \cdot 3a \cdot \dots \cdot (p-1)a \equiv (p-1)! \pmod{p}.$$

Since $(p-1)! \not\equiv 0 \pmod{p}$, we can divide both sides by it and get:

$$a^{p-1} \equiv 1 \pmod{p}.$$

We can conclude that $a^p \equiv a \pmod{p}$ holds \square .

Witness

Fix an integer n . We say that an integer a is a **Fermat witness** for the compositeness of n if and only if

$$a^n \not\equiv a \pmod{n}$$

holds.

Example

Bob wants to know whether

$$n = 456\,989\,977\,669$$

is a prime number.

The answer is a resounding **no**, since

$$2^n \equiv 1\,493\,546 \not\equiv 2 \pmod{n},$$

so 2 is a Fermat witness for the compositeness of n .

Note that we did not need to find the factorization of n to establish the compositeness. We have

$$n = 456\,989\,977\,669 = p_{50\,000}p_{60\,000}$$

so

$$456\,989\,977\,669 = 611\,953 \times 746\,773.$$

Often trying a small number of potential witnesses will reveal much quicker that n is composite than factorization.

Lemma

Let n be odd. If n has a Fermat witness b for compositeness, then at least half of the elements in

$$\mathbf{Z}_n^* = \{a \in \mathbf{Z}_n \mid \gcd(a, n) = 1\}$$

are Fermat witnesses for compositeness.

Let $S = \{a \in \mathbf{Z}_n^* \mid a^{n-1} \equiv 1 \pmod{n}\}$ be the set of non-witnesses. Then $\{ab \mid a \in S\}$ is a set of $|S|$ distinct witnesses, since $(ab)^{n-1} \not\equiv 1 \pmod{n}$.

Fermat Test Algorithm

Input: a positive integer $n \geq 2$.

for $i = 1$ **to** t **do**

Choose an integer a in the range $2 \leq a < n$ uniformly at random.

return 'composite' **if** $a^n \not\equiv a \pmod{n}$.

od;

return 'potentially prime'

Consider $n = 561$.

Then

$$a^{561} \equiv a \pmod{561}$$

for all a in the range $1 \leq a \leq 560$. But

$$561 = 3 \times 11 \times 17.$$

Nasty numbers such as 561 that have no Fermat witnesses of compositeness are called Carmichael numbers.

Fermat Test

The Fermat test **cannot prove primality** with certainty.

However, it **can prove compositeness**.

Flaw

The Fermat test systematically fails to detect that Carmichael numbers are composite.

There exist an infinite number of Carmichael numbers.

Conclusion

We need a better concept for the witnesses.

Miller-Rabin

Proposition

Let p be an odd prime and write

$$p - 1 = 2^k q \quad \text{with integers } k \geq 0 \text{ and } q \text{ odd.}$$

Let a be any positive integer not divisible by p . Then one of the following conditions is true:

- (a) $a^q \equiv 1 \pmod{p}$.
- (b) One of $a^q, a^{2q}, a^{4q}, \dots, a^{2^{k-2}q}, a^{2^{k-1}q}$ is $\equiv -1 \pmod{p}$.

Proof of the Proposition

By Fermat's Little Theorem, we have $a^{q-1} \equiv a^{2^k q} \equiv 1 \pmod{p}$.

Thus, in the list

$$a^q, a^{2q}, a^{4q}, \dots, a^{2^{k-1}q}, a^{2^k q}$$

the last one is congruent to 1 and each number is the square of the previous number. Then we either have

- 1 the first number satisfies $a^q \equiv 1 \pmod{p}$,
- 2 there must be some number in the list that is $\not\equiv 1 \pmod{p}$.

If a number b satisfies

$$b \not\equiv 1 \pmod{p} \quad \text{and} \quad b^2 \equiv 1 \pmod{p},$$

then it satisfies $b \equiv -1 \pmod{p}$.

Miller-Rabin Witness

Let n be an odd positive integer and write $n - 1 = 2^k q$ with q odd. An integer a satisfying $\gcd(a, n) = 1$ is called a **Miller-Rabin witness** for n if and only if the following two conditions are satisfied:

- 1 $a^q \not\equiv 1 \pmod{n}$,
- 2 $a^{2^j q} \not\equiv -1 \pmod{n}$ for all $j = 0, 1, 2, \dots, k - 1$.

Example

Consider the Carmichael number $n = 561$.

Then $n - 1 = 560 = 2^4 \cdot 35$.

For $a = 2$ and $q = 35$, we get

$$2^{35} \equiv 263 \not\equiv 1 \pmod{561}$$

and

$$2^{35} \equiv 263 \not\equiv -1 \pmod{561}$$

$$2^{70} \equiv 166 \not\equiv -1 \pmod{561}$$

$$2^{140} \equiv 67 \not\equiv -1 \pmod{561}$$

$$2^{280} \equiv 1 \not\equiv -1 \pmod{561}$$

Thus, 2 is a Miller-Rabin witness for compositeness of $n = 561$.

Proposition

Let n be an odd composite number. Then at least 75% of the numbers a between 1 and $n - 1$ are Miller-Rabin witnesses for n .

Miller-Rabin Primality Test

Input: a positive integer $n \geq 2$.

for $i = 1$ **to** t **do**

Choose an integer a in the range $2 \leq a < n$ uniformly at random.

return *'composite'* **if** a is MR-Witness (mod n).

od;

return *'prime'* // potentially incorrect

We want to know

$$\Pr[n \text{ prime} \mid \text{'prime'}] = ?$$

We do know

$$\Pr[\text{'composite'} \mid n \text{ prime}] = 0$$

$$\Pr[\text{'prime'} \mid n \text{ prime}] = 1$$

$$\Pr[n \text{ composite} \mid \text{'composite'}] = 1$$

$$\Pr[\text{'prime'} \mid n \text{ composite}] = \left(\frac{1}{4}\right)^t$$

$$\frac{\pi(n)}{n} \sim \frac{1}{\ln n}.$$

Thus, for large n , we have

$$\Pr[n \text{ is prime}] \approx \frac{1}{\ln n} \quad \text{and} \quad \Pr[n \text{ is composite}] \approx \frac{\ln n - 1}{\ln n}$$

$$\Pr[A|B] = \frac{\Pr[B|A] \Pr[A]}{\Pr[B|A] \Pr[A] + \Pr[B|\bar{A}] \Pr[\bar{A}]}$$

Probability that Miller Rabin Correctly Identifies a Prime (1/3)

$$\Pr[A|B] = \frac{\Pr[B|A] \Pr[A]}{\Pr[B|A] \Pr[A] + \Pr[B|\bar{A}] \Pr[\bar{A}]}$$

Let $A = n$ prime, and $B = \text{'PRIME'}$.

$$\Pr[n \text{ prime} | \text{'PRIME'}] =$$

$$\frac{\Pr[\text{'PRIME'} | n \text{ prime}] \Pr[n \text{ prime}]}{\Pr[\text{'PRIME'} | n \text{ prime}] \Pr[n \text{ prime}] + \Pr[\text{'PRIME'} | n \text{ composite}] \Pr[n \text{ composite}]}$$

Probability that Miller Rabin Correctly Identifies a Prime (2/3)

$$\Pr[n \text{ prime} | \text{'PRIME'}] =$$

$$\frac{\Pr[\text{'PRIME'} | n \text{ prime}] \Pr[n \text{ prime}]}{\Pr[\text{'PRIME'} | n \text{ prime}] \Pr[n \text{ prime}] + \Pr[\text{'PRIME'} | n \text{ composite}] \Pr[n \text{ composite}]}$$

$$\Pr[n \text{ prime} | \text{'PRIME'}] = \frac{1 \cdot (1/\ln n)}{1 \cdot (1/\ln n) + \frac{1}{4^t} (\ln n - 1)/\ln n}$$

$$\Pr[n \text{ prime} | \text{'PRIME'}] = \frac{1}{1 + \frac{1}{4^t}(\ln n - 1)}$$

Thus, if $t \geq \log_4(\ln n - 1)$, then $\Pr[n \text{ prime} | \text{'PRIME'}] \geq 1/2$.

If $t = 5$, then you can determine with probability $1/2$ or greater whether a 1024 bit number is prime.

Primality tests were among the first randomized algorithms.

The Miller-Rabin primality test is an example of a Monte-Carlo randomized algorithm with one-sided error (it never errs when declaring 'composite', but it might err when declaring 'prime').

With a few repetitions, we can keep the probability of error very low. If you choose $t \geq 30$ repetitions, then the chance that your computer hardware will make a mistake in the calculations is more likely than that the probability test fails.

We will need a few facts from probability theory, but not too many!
We will review everything that we will need.