Primality Tests

Andreas Klappenecker

Texas A&M University

Question

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.

AKS Primality Test

- Use the Agrawal-Kayal-Saxena primality test.
- It is a deterministic $\widetilde{O}(\log(n)^{12})$ 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

Goal

We will now develop some randomized algorithms for primality testing.

Notation: Congruence Relations

Congruence Relation

Congruence Relation

Let a, b, and n be integers. We write

$$a \equiv b \pmod{n}$$

if and only if the integer $a - b$ is divisible by n.

Example

- $39 \equiv 9 \pmod{15}$
- $1001 \equiv 1 \pmod{10}$
- $a \equiv b \pmod{n}$ means that a and b have the same remainder when divided by n.

Important Properties

Congruence Properties

If $a_1 \equiv a_2 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$, then

- $a_1 + b_1 \equiv a_2 + b_2 \pmod{n}$,
- $a_1 - b_1 \equiv a_2 - b_2 \pmod{n}$,
- $a_1 \cdot b_1 \equiv a_2 \cdot b_2 \pmod{n}$.

These properties follow easily from the definition.

What about division?

Bezout's Theorem

If a and b are integers, and $g = \gcd(a, b)$, then there exist integers a' and b' such that

$$\gcd(a, b) = aa' + bb'.$$

This follows from the Euclidean algorithm. Recall that this algorithm performs successive quotient/remainder calculations, $a = bq_1 + r_1$, replaces a and b by b and r_1, and repeats until the remainder is 0. In matrix notation,

$$\underbrace{\begin{pmatrix} 0 & 1 \\ 1 & -q_k \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & -q_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -q_1 \end{pmatrix}}_{= \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \gcd(a, b) \\ 0 \end{pmatrix}$$

Important Properties

If $\gcd(a, n) = 1$, then there exists an integer a' such that

$$aa' \equiv 1 \pmod{n}.$$

In other words, a has then a multiplicative inverse.

Indeed, by Bezout's theorem,

$$\gcd(a, n) = 1 = aa' + nn'$$

for some integers a' and n'. Reducing this equation modulo n, we get

$$1 \equiv aa' \pmod{n}.$$

We usually denote the multiplicative inverse of a by a^{-1}, so a^{-1} is the integer a'.

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.

Proof of Fermat's Little Theorem (1/2)

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

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

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 $1 \leqslant j, k < p$.

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

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

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

Multiplying these numbers together, we can conclude that

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

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.

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.

Example

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 $n \geqslant 2$.

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

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

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

od;

return 'potentially prime'

Snag

Consider $n = 561$.
Then

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

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

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

Nasty numbers such as 561 that have no Fermat witnesses of compositeness are called Carmichael numbers.
[In other words, a **Charmichael number** is a composite number n such that $b^n \equiv b \pmod{n}$ holds for all integers b.]

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

Let p be a prime and x an integer such that

$$x^2 \equiv 1 \pmod{p}.$$

Then $x^2 - 1$ is a difference of squares, and we get

$$(x-1)(x+1) \equiv 0 \pmod{p}.$$

Therefore, we can conclude that either

$$x \equiv 1 \pmod{p} \quad \text{or} \quad x \equiv -1 \pmod{p}.$$

Observation

If p is an odd prime and a is an integer not divisible by p, then

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

by Fermat's little theorem. Since $p-1$ is even and $a^{p-1} \equiv 1 \pmod{p}$, we have

$$a^{(p-1)/2} \equiv \pm 1 \pmod{p}.$$

This is another condition that we can check.

If $(p-1)/2$ is even and $a^{(p-1)/2} \equiv 1 \pmod{p}$, then

$$a^{(p-1)/4} \equiv \pm 1 \pmod{p}.$$

We can continue in this fashion.

Proposition

Let p be an odd prime and write

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

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 $a^q, a^{2q}, a^{4q}, \ldots, a^{2^{k-2}q}, a^{2^{k-1}q}$ is $\equiv -1 \pmod{p}$.

Proof of the Proposition

By Fermat's Little Theorem, we have $a^{p-1} \equiv a^{2^k q} \equiv 1 \pmod{p}$. 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 b in the list such that $b \not\equiv 1 \pmod{p}$ and $b^2 \equiv 1 \pmod{n}$. A integer b satisfying

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

must satisfy $b \equiv -1 \pmod{p}$.

Recap

Let us reiterate

If p be an odd prime,

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

and a is a positive integer not divisible by p, then one of the following conditions is true:

(a) $a^q \equiv 1 \pmod{p}$.

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

What is the negation of this statement?

Miller-Rabin Witness

Let n be an odd positive integer and write $n-1=2^kq$ 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 $j=0,1,2,\ldots,k-1$.

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

$$\begin{aligned}
2^{35} &\equiv 263 \not\equiv -1 \pmod{561} \\
2^{70} &\equiv 166 \not\equiv -1 \pmod{561} \\
2^{140} &\equiv 67 \not\equiv -1 \pmod{561} \\
2^{280} &\equiv 1 \not\equiv -1 \pmod{561}
\end{aligned}$$

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

Abundance of Miller-Rabin Witnesses

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 $n \geqslant 2$.

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

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

return *'composite'* **if** a *is MR-Witness* $\pmod{n}$.

od;

return 'prime' // potentially incorrect

Probabilities

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

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

Bayes Formula

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

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

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

Let $A = n$ prime, and $B =$ '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)

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

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

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

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

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

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

Conclusions

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 $t \geqslant 30$ repetitions, then the chance that your computer hardware will make a mistake in the calulations 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.