High School Student Institute of Mathematics and Informatics

Summer Research School Symposium 2021

Primality Testing

Author:

Emilie Ma
University of British Columbia
kewbish@gmail.com

Scientific Advisor:

Pressiana Marinova
Occado Technology Sofia
pressiana.marinova@gmail.com

Abstract

Abstract

1 Introduction

Introduction

2 Background Theory

2.1 Fermat Primality Test

The Fermat Primality Test is a probablistic primality test based on Fermat's little theorem. Fermat's little theorem, developed by Pierre de Fermat in 1640, states that for any integer a and any prime p, the following holds:

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

If a is not divisible by, or *coprime* to, p, the following is equivalent:

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

Proof. Consider $S = \{a, 2a, 3a, \ldots, (p-1) * a\}$. Suppose ra and sa in the set are equal $\pmod{p}$, so $r \equiv s \pmod{p}$. Therefore, the $p-1$ multiples of a in S are uniquely distinct, and must be congruent to $1, 2, 3, \ldots, (p-1)$ in some order. Multiply these congruences like so:

$$a * 2a * 3a * \ldots * (p-1)a \equiv 1 * 2 * 3 * \ldots (p-1) \pmod{p}$$

This gives:

$$a^{(p-1)} * (p-1)! \equiv (p-1)! \pmod{p}$$

Divide by $(p-1)!$ on each side for:

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

To arrive at the alternate form of Fermat's Little Theorem, multiply both sides by a.

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

□

Knowing that $a^{(n-1)} \equiv 1 \pmod{n}$ holds if n is prime, Fermat's primality test chooses k random integers a coprime to n to test if all a are congruent to 1. Because this holds trivially for $a \equiv 1 \pmod{n}$ and if n is odd and $a \equiv -1 \pmod{n}$, a is conventionally chosen such that $1 < a < n-1$. Higher values of k indicate a higher probability that the number is prime.

If n passes these k base tests, it is known as a probable prime. However, not all numbers that pass the Fermat primality test are prime - composite numbers n that pass the test are known as Fermat pseudoprimes. There are infinitely many Fermat pseudoprimes, and several forms of composite numbers that pass the test. For example, Carmichael numbers, composite numbers that satisfy the relation $b^{(n-1)} \equiv 1 \pmod{n}$ for all integers b coprime to n, all pass Fermat's primality test.

2.2 Euler (Solovay-Strassen) Test

The Solovay-Strassen Test is another probablistic test, utilizing the properties of Euler's theorem. Proposed by Leonhard Euler in 1763, Euler's theorem is a generalization of Fermat's little theorem, stating that if a and p are coprime, then the following holds:

$$a^{\phi(n)} \equiv 1 \pmod{n}$$

The function $\phi(n)$ is Euler's totient function. The totient of some number n is the number of positive integers l in the range $1 <= l <= n$ where l is coprime to n.

Proof. Consider $S = \{1 <= l <= n | gcd(l, n) = 1\} = \{l_1, l_2, l_3, \ldots, l_{\phi(n)}\}$. Create a set $aS = \{al_1, al_2, al_3, \ldots, al_{\phi(n)}\}$.
All elements of aS are relatively prime to n, so if all elements of aS are distinct, $aS = S$. All elements of aS are distinct, as all elements of S are distinct. Therefore, each element of $aS \equiv S \pmod{n}$. Therefore:

$$l_1 * l_2 * l_3 * \ldots * l_{\phi(n)} \equiv al_1 * al_2 * al_3 * \ldots * al_{\phi(n)} \pmod{n}$$

As $l_1 * l_2 * l_3 * \ldots * l_{\phi(n)}$ is relatively prime to n, reducing this gives:

$$a^{\phi(n)} * l_1 * l_2 * l_3 * \ldots * l_{\phi(n)} \equiv l_1 * l_2 * l_3 * \ldots * l_{\phi(n)} \pmod{n}$$

Therefore, $a^{\phi(n)} \equiv 1 \pmod{n}$. □

Fermat's little theorem is considered a special case of Euler's theorem, because if n is prime, $\phi(n) = n - 1$.

Given that $a^{\phi(n)} \equiv 1 \pmod{n}$, then

$$a^{\phi(n)/2} \equiv \begin{cases} 1 \pmod{n} & \text{when there exists } x \text{ such that } a \equiv x^2 \pmod{n} \\ -1 \pmod{n} & \text{when there is no such integer.} \end{cases}$$

The conditions above form the criteria for the Legendre symbol of a and n. The Legendre symbol $(\frac{a}{n})$ is defined like so:

$$(\frac{a}{n}) \begin{cases} 0 & \text{when } a \equiv 0 \pmod{n} \\ -1 & \text{when } a \not\equiv 0 \pmod{n} \text{ and there exists } x : a \equiv x^2 \pmod{n} \\ -1 & \text{when } a \not\equiv 0 \pmod{n} \text{ and there is no such integer } x. \end{cases}$$

The Jacobi symbol is the generalization of the Legendre symbol to any odd integer n, and is used in the Solovay-Strassen primality test. It is defined as the product of the Legendre symbols of n's prime factors, such that:

$$(\frac{a}{n}) = (\frac{a}{p_1})^{\alpha_1} * (\frac{a}{p_2})^{\alpha_2} * \ldots * (\frac{a}{p_k})^{\alpha_k}$$

for $n = p_1^{\alpha_1} * p_2^{\alpha_2} * \ldots * p_k^{\alpha_k}$.

As with Fermat's primality test, k random bases a are tested. If $a^{(n-1)/2} \equiv (\frac{a}{n}) \pmod{n}$ holds for all k basis, then n is a probable prime.

Similar to Fermat's primality test, the Solovay-Strassen test may pass composite numbers as primes. These are then known as Euler (sometimes Euler-Jacobi) pseudoprimes or liars. All Euler pseudoprimes are also Fermat pseudoprimes.

2.3 Miller-Rabin Primality Test

A third probablistic primality test, the Miller-Rabin primality test was discovered first by Gary Miller in 1976, and subsequently modified by Michael Rabin in 1980. The test relies on two congruence relations that hold when n is an odd prime and rewritten as $2^s * d + 1$, and a is a base such that $0 < a < n$:

$$a^d \equiv 1 \pmod{n}$$
$$a^{(2^r * d)} \equiv -1 \text{ for some } r \text{ such that } 0 <= r < s$$

Because n is written as $2^s * d + 1$, $n - 1 = 2^s * d$. Therefore, if $a^d \equiv \pm 1 \pmod{n}$, then n is a strong probable prime.

Proof. Given that:

$$a^{(n-1)} \equiv (a^d)^{2^s} \equiv 1 \pmod{n}$$

for all prime n, and because there are no square roots of 1 other than ± 1, the repeated squaring with 2^s doesn't affect the congruence. □

Otherwise, $a^d \pmod n$ is squared, for a^2d. If $a^2d \equiv 1 \pmod n$, n is composite, because there are different square roots of $a^2m \pmod n$ other than ± 1. If $a^2d \equiv -1 \pmod n$, then n is a probable prime for similar reasons as above.

These checks are repeated until $a^{(2^{(s-1)}*d)}$ has been reached. If it is ± 1, the result is known by the tests above; however, if not, n is composite, by Fermat's little theorem.

3 Agarwal-Kayal-Saxena Primality Test

By contrast, the Agarwal-Kayal-Saxena, or AKS, primality test, is a deterministic primality test first proprosed by Manindra Agarwal, Neeraj Kayal, and Nitin Saxena in 2002. It is the first primality test to deterministically verify primality in polynomial time for all number inputs, with a time complexity bound of $\mathcal{O}(\log(n)^{21/2})$.

The test is based on the theorem that an integer $n >= 2$ and an integer a coprime to n, n is prime if and only if the below holds within the polynomial ring $\mathbb{Z}[x]$, or the ring of polynomials with degree at most n over $\mathbb{Z}$.

$$(X + a)^n \equiv X^n + a \pmod n$$

Proof. This is a generalization of Fermat's little theorem over polynomials, proven with the binomial identity.

$$(x + a)^n = \sum_{i=0}^{n} \binom{n}{i} x^i a^{n-i} where \binom{n}{0} = \binom{n}{n} = 1$$

If n is prime, then:

$$\binom{n}{i} = \frac{n(n-1)\dots n-i+1}{i!}$$

for all i > 1. When taken $\pmod n$, n divides the numerator once; therefore, $\binom{n}{i} \equiv 0 \pmod n$. a does not necessarily need to be coprime with n. If n is not prime, and a is coprime with n, then n has a prime factor of form p^k. p^k will divide n, but p^{k-1} will not. Given the monomial with a coefficient of:

$$\binom{n}{i} = \frac{n(n-1)\dots n-p+1}{p!}$$

p^k divides n, but not $(n-1)\dots(n-p+1)$. Therefore, $p^k | n(n-1)\dots(n-p+1)$ but $p^{k+1} \nmid n(n-1)\dots(n-p+1)$. p also divdes p, but not $1\dots p$, so $p|p!$ but $p^2 \nmid p!$. This gives $p^{k-1} | \binom{n}{p}$ and $p^k \nmid \binom{n}{p}$. As p^k is a factor of n, $n \nmid \binom{n}{p}$, and x^p does not vanish. □

The test begins by ensuring n is not a perfect power, or a number of form a^k for some a and k. If it is, the number is composite.

Next, the algorithm finds the smallest r such that $ord_r(n) > \log 2n^2$ and r is coprime to n.

The algorithm then checks for all $2 <= a <= min(r, n-1)$ that $a \nmid n$. If $a|n$ for some a in this range, the number is composite. This is equivalent to trial division up to r.

If $n <= r$, the number is prime, as this is equivalent to trial division to $\sqrt{n}$.

The last step of the algorithm reduces the time complexity from exponential to polynomial time by operating over the finite ring $(\mathbb{Z}/(n))[X]/(X^r - 1)$. Each a from $1 <= a <= \lfloor \sqrt{\varphi(r)} \log_2(n) \rfloor$ is checked for the generalized Fermat's theorem. If it does not pass, the number if composite; if it passes, the number is prime.

4 Methods

Methods

4.1 Base Analysis

Each primality test was analyzed in a standard way over three trials; the raw data is available in Appendix A. Each trial tested a different k value, and consisted of:

- Generating a random set (S) of 10^4 integers such that $10^6 < x < 2 * 10^6$
- Using SageMath's `is_prime` to check for primality for each integer in S
- Running the primality test with k attempts run on each integer in S with respect to some base a
- Counting all pseudoprimes which passed the primality test but not Sage's primality test
- Repeat for three sub-trials, average results and return lowest number of bases tried (lowest k) that returned the lowest number of pseudoprimes passed

a was a choice of either all random bases, 2, 3, 5, and a pair of 2, 3, or 5. Each trial was timed with the Linux `time` command, recording the real, or total wall time, elapsed.

5 Results

Simplified tables and synthesized figures

The results over three trials of the primality tests are shown in Tables 1, 2, 3, 4 in Appendix A.

6 Discussion

Discussion of results

7 Conclusion

Conclusion

References

Appendix A Raw Data for Primality Tests

A.1 Fermat's Primality Test

Table 1: Raw data for Fermat's Primality Test Trials

	All Random Bases		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.934s	0m9.884s	0m10.670s
Lowest k required	2	2	3
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.0606	0.08080	0.0909
Range of lowest k required	1		
Range of number of pseudoprimes passed	0		
	Base 2		
	Trial 1	Trial 2	Trial 3
Running Time	0m6.881s	0m6.233s	0m6.487s
Lowest k required	61	14	47
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	3.8989	3.9090	4.0303

Range of lowest k required	47		
Range of number of pseudoprimes passed	0		
	Base 3		
	Trial 1	Trial 2	Trial 3
Running Time	0m6.710s	0m5.965s	0m6.912s
Lowest k required	21	6	52
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	3.3434	3.9292	3.4040
Range of lowest k required	46		
Range of number of pseudoprimes passed	0		
	Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m7.535s	0m6.277s	0m6.064s
Lowest k required	25	19	41
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	3.4646	3.5858	3.7474
Range of lowest k required	22		
Range of number of pseudoprimes passed	0		
	Base 2 and Base 3		
	Trial 1	Trial 2	Trial 3
Running Time	0m9.046s	0m8.423s	0m7.899s
Lowest k required	2	5	2
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.8787	0.8282	0.8888
Range of lowest k required	3		
Range of number of pseudoprimes passed	0		
	Base 3 and Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.304s	0m8.838s	0m7.616s
Lowest k required	4	1	2
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.7474	0.8888	0.7171
Range of lowest k required	3		
Range of number of pseudoprimes passed	0		
	Base 2 and Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m10.217s	0m8.388s	0m8.096s
Lowest k required	2	11	2
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.7777	0.9696	0.6868
Range of lowest k required	9		
Range of number of pseudoprimes passed	0		

A.2 Euler's (Solovay-Strassen) Primality Test

Table 2: Raw data for Euler (Solovay-Strassen) Primality Test Trials

	All Random Bases		
	Trial 1	Trial 2	Trial 3
Running Time	0m4.859s	0m5.044s	0m4.888s

Lowest k required	3	2	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.0202	0.0404	0.0000
Range of lowest k required	2		
Range of number of pseudoprimes passed	0		

	Base 2		
	Trial 1	Trial 2	Trial 3
Running Time	0m5.026s	0m5.925s	0m5.145s
Lowest k required	1	1	4
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	1.3333	1.6262	1.4747
Range of lowest k required	3		
Range of number of pseudoprimes passed	0		

	Base 3		
	Trial 1	Trial 2	Trial 3
Running Time	0m6.710s	0m6.443s	0m4.905s
Lowest k required	3	3	3
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	1.6464	1.5050	1.3737
Range of lowest k required	0		
Range of number of pseudoprimes passed	0		

	Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m5.031s	0m6.367s	0m4.978s
Lowest k required	6	2	7
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	1.7979	1.4242	1.3535
Range of lowest k required	5		
Range of number of pseudoprimes passed	0		

	Base 2 and Base 3		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.009s	0m7.581s	0m7.025s
Lowest k required	1	1	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.3030	0.4949	0.2626
Range of lowest k required	0		
Range of number of pseudoprimes passed	0		

	Base 3 and Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.597s	0m6.908s	0m6.757s
Lowest k required	1	2	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.1414	0.1818	0.2222
Range of lowest k required	1		
Range of number of pseudoprimes passed	0		

	Base 2 and Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.354s	0m6.952s	0m6.849s
Lowest k required	1	1	2
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.2323	0.2222	0.3333

Range of lowest k required	1
Range of number of pseudoprimes passed	0

A.3 Miller-Rabin's Primality Test

Table 3: Raw data for Miller-Rabin Primality Test Trials

	All Random Bases		
	Trial 1	Trial 2	Trial 3
Running Time	0m7.274s	0m7.391s	0m7.806s
Lowest k required	1	2	2
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.0000	0.0101	0.0202
Range of lowest k required	1		
Range of number of pseudoprimes passed	0		
	Base 2		
	Trial 1	Trial 2	Trial 3
Running Time	0m5.780s	0m6.143s	0m6.181s
Lowest k required	2	1	3
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	1.2222	1.0101	1.0303
Range of lowest k required	2		
Range of number of pseudoprimes passed	0		
	Base 3		
	Trial 1	Trial 2	Trial 3
Running Time	0m5.651s	0m5.768s	0m7.215s
Lowest k required	1	4	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	1.0808	0.8080	0.7575
Range of lowest k required	3		
Range of number of pseudoprimes passed	0		
	Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m5.676s	0m5.707s	0m6.111s
Lowest k required	1	12	13
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	1.0808	1.1919	1.2727
Range of lowest k required	12		
Range of number of pseudoprimes passed	0		
	Base 2 and Base 3		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.059s	0m8.315s	0m8.202s
Lowest k required	1	1	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.0303	0.0808	0.0808
Range of lowest k required	0		
Range of number of pseudoprimes passed	0		
	Base 3 and Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.132s	0m8.172s	0m8.314s

Lowest k required	1	1	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.1111	0.1111	0.1010
Range of lowest k required	0		
Range of number of pseudoprimes passed	0		
	Base 2 and Base 5		
	Trial 1	Trial 2	Trial 3
Running Time	0m8.321s	0m8.463s	0m8.752s
Lowest k required	1	1	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.0505	0.0606	0.0808
Range of lowest k required	0		
Range of number of pseudoprimes passed	0		

A.4 Probablistic AKS Primality Test

Table 4: Raw data for Probablistic AKS Primality Test Trials

	1 Equation Tested		
	Trial 1	Trial 2	Trial 3
Running Time	1m22.615s	1m30.450s	1m22.402s
Lowest k required	1	1	1
Pseudoprimes passed at lowest k	0	0	0
Average pseudoprimes passed	0.0000	0.0000	0.0000
Range of lowest k required	0		
Range of number of pseudoprimes passed	0		