Mathematicians have tried in vain to this day
to discover some order in the sequence of prime numbers,
and we have reason to believe that it is a mystery into
which the human mind will never penetrate.
L. EULER [**9**]

Methods of Primality Testing

ZACHARY S. MCGREGOR-DORSEY

Abstract. We discuss the most popular methods of primality testing, along with some intermediate steps of their formulation. First, we review the importance of primality testing, the history of prime numbers, and the difficulties implementing the tests. Then we present the Chinese Primality Test, Fermat's Little Theorem, and the Strong Psuedoprimality Test. Finally, we discuss Lucas Sequences and the Lucas Test.

1. Introduction. The interest in primality testing has grown rapidly in the past two decades since the introduction of public-key cryptography, now the standard form of encryption for electronic correspondence. The security of this type of cryptography primarily relies on the difficulty involved in factoring very large numbers. Integer factorization poses many problems, a key one being the testing of numbers for primality. A reliable and fast test for primality would bring us one step closer to decoding billions of bits of highly confidential information. Therefore, the mathematics and computer science communities have begun to address the problem of primality testing with increased vigor. A primality test is simply a function that determines if a given integer greater than 1 is prime or composite. This paper addresses some attempts at developing an efficient and reliable method for testing primality.

Despite the recent interest in primality testing, the quest for discovering a good test is by no means a new one, and very likely one of the oldest issues in mathematics. The ancient priests in Uruk the Sheepfold, circa 2500 B.C., are known to have inscribed long lists of prime numbers in cuneiform. Both the ancient Greeks and the ancient Chinese independently developed primality tests [**10**, p. 47]. One of the simplest and most famous primality test is the Sieve of Eratosthenes, developed one millenium later than the Chinese method.

Eratosthenes lived in Greece circa 200 B.C. His method for determining primality is as follows. Suppose we want to determine if n is prime. First, we make a list of all integers $2, 3, \dots, m$ where m is the largest integer less than or equal to $\sqrt{n}$. Next, we circle 2 and cross off from the list all multiples of two. Then we circle 3 and cross off its multiples. We now continue through the list, each time advancing to the least integer that is not crossed off, circling that integer, and crossing off all its multiples. We then test to see if any of the circled numbers divide n. If the list of circled numbers is exhausted and no divisor is found, then n is prime. This algorithm is based on the

simple observation that, if n is composite, then n has a prime factor less than or equal to $\sqrt{n}$.

The Sieve of Eratosthenes exemplifies both the simplicity of testing for primality and restraints on the efficiency of such tests. The algorithm itself is a fairly straightforward process and easy to implement, based almost completely on the definition of primes. However, though the algorithm is easy to implement, it is by no means efficient.

As in the application to cryptography, most primality testing is concerned with large numbers, usually in excess of 100 digits and often much larger. If we were to use the Sieve of Eratosthenes to determine the primality of a number with just 20 digits, we would need first to find at least all the primes up to 10^{10}. According to [10, p. 47], there are around 450 million primes less that 10^{10}. At the rate of finding one prime per second (including crossing off all of its multiples!), we would be working for a little over 14 years to find the 450 million primes, which then would have to be divided into our original 20-digit number. Several decades is a fairly significant time period to determine if just one number—a relatively small number—is prime. Certainly there are computers that can run this algorithm faster and more efficiently by, for example, keeping all primes in a database to reduce redundant testing. Still, the overall inefficiency of the Sieve of Eratosthenes makes it impractical to use; we need quicker algorithms.

We can approach the problem of finding a faster algorithm for primality testing in several ways. One way is to find a pattern among primes, and then determine if a given integer n follows that pattern. So far, no exhaustive and easily implementable pattern has been found. Another method is to find a pattern among all composites, and then determine if a given n follows that pattern. In essence, these two methods are the same, since if an integer is not prime, then it is composite and vice versa.

The approaches discussed below are all based on finding patterns that are unique to composites. However, these approaches are not perfect. The patterns that they use fit most, but not all, composites. In other words, the numbers fitting the patterns are always composite, but there are some composites that do not fit the pattern. Composites that do not fit the patterns are hereafter called *pseudoprimes*. Despite the pseudoprime shortcoming, tests with almost perfect accuracy are quite useful in many applications.

Some tests, such as the Sieve of Eratosthenes, fully establish primality, but do so much faster. Though not discussed in this paper, it is important to note the existence of such tests. One such test is the Elliptic Curve Primality Procedure (ECPP). However, whereas composite-based tests can digest a 500-digit number in only a few minutes, ECPP takes several hours [10, p. 71]. Hence, though ECPP is much more efficient than the Sieve of Eratosthenes, it is not nearly as efficient as other tests. Generally, ECPP is used to verify results given by the other tests.

Section 2 describes the Chinese primality test. Section 3 describes a generalization of this test, discovered by P. de Fermat. Section 4 describes the final step in this series called the Strong Pseudoprimality Test, also refered to as the Miller–Rabin Primality Test. Section 5 describes a different approach with the Lucas Primality Test. All these tests can be performed in polynomial time or better.

2. Chinese Primality Test. The Sieve of Eratosthenes, though accurate, is very tedious and inefficient. We can improve on its efficiency by looking for characteristics that some composite numbers share as well. In approximately 500 B.C., according to [3, p. 59], the ancient Chinese discovered that, if p is prime, then $2^p - 2$ is divisible by p. We state this fact in its contrapositive form in the following proposition.

Proposition 2-1. *Let n be a positive integer greater than 1. If 2^n is not congruent to 2 (mod n), then n is a composite.*

Proof: The proof is left for Section 3, where a more general case is discussed. □

Note that this proposition says nothing directly about primes. We write so, because although for any prime p, the difference $2^p - 2$ is divisible by p, nevertheless there are also some composite numbers that pass the test as well. In fact, there are infinitely many of such composite numbers. We now give a name to these composites.

Definition 2-2. Let n be a composite number. If n divides $2^n - 2$, then n is called a *base-2 pseudoprime*.

There are very few such pseudoprimes. In fact, a number that passes the Chinese Primality Test has only a 0.002% chance of not being prime. Below 200, there are only six such pseudoprimes: 341, 561, 645, 1105, 1729, and 1905. Base-2 pseudoprimes are evidently few and far between. If we define $\pi(n)$ to be the number or primes less than n, and $\pi'(n)$ to be the number of base-2 pseudoprimes less than n, we find that the ratio $\pi(n) : \pi'(n)$ generally increases as n increases.

We can now state the Chinese Primality Theorem as the converse to Proposition 2-1.

Theorem 2-3 (Chinese Primality Theorem). *Let n be an integer, $n > 1$. If 2^n is congruent to 2 (mod n), then n is either a prime or a base-2 pseudoprime.*

Proof: The assertion follows directly from Proposition 2-1 and from the definition of a base-2 pseudoprime. □

3. Fermat Primality Test. In 1640, Fermat rediscovered what the ancient Chinese had known nearly 2000 years before him. He also examined the problem using bases other than 2, improving on the accuracy of the Chinese test. The result of his work is now known as Fermat's Little Theorem. As was characteristic of Fermat, he never gave a proof for this theorem, saying "I would send you the demonstration, if I did not fear its being too long," [**2**, p. 79]. The theorem was first proved by Leibniz about forty years later. Here is the precise statement.

Theorem 3-1 (Fermat's Little Theorem). *Let p be a prime, and a any positive integer. If $\gcd(p, a) = 1$, then*

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

Proof: Consider $a, 2a, \dots, (p-1)a$, the first $p-1$ positive multiples of a. Take two such multiples, ma and na. If they are congruent modulo p, then we have $m \equiv n$ (mod p), because p does not divide a. Hence, the $p-1$ multiples of a are distinct and nonzero. Thus each multiple must be congruent to a different member of the set $\{1, 2, \dots, p-1\}$. Because congruence preserves multiplication, we have the following congruence:

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

Rearranging the terms gives us

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

Dividing $(p-1)!$ into both sides completes the proof. □

In Fermat's Little Theorem, if we replace a with 2, and multiply both sides by 2, we then have the Chinese Primality Theorem.

Fermat's Little Theorem, because it gives us a property of all primes, works wonderfully as the basis for many primality tests, including Section 4's Strong Pseudoprimality Test. For this application, we use its contrapositive:

Theorem 3-2 (Fermat's Primality Test). *An odd positive integer n is composite if there exists a positive integer a such that*

$$\gcd(a, n) = 1 \ and \ a^{n-1} \not\equiv 1 \pmod{n}.$$

The contrapositive is used because, like the Chinese Primality Test, Fermat's Little Theorem is troubled by pseudoprimes. For any given a, a composite number n such that $\gcd(a, n) = 1$ and $a^{n-1} \equiv 1 \pmod{n}$ is called a *base-a pseudoprime*. With the Chinese Test, we saw that base-2 pseudoprimes are fairly infrequent. Other bases have a similar low frequency of pseudoprimes. More important, different bases result in mostly different pseudoprimes. Therefore Fermat's Little Theorem is much better than the Chinese Primality Theorem for determining primality: If a number passes the test for base 2, we can eliminate error by checking our results with more bases. Unfortunately, some composites are pseudoprimes for all bases.

Definition 3-3. Let n be a composite. If $a^{n-1} \equiv 1 \pmod{n}$ for every positive integer a with $\gcd(a, n) = 1$, then n is called a *Carmichael number*.

As one would expect, Carmichael numbers are few and far between. Richard Pinch (unpublished) has recently found that there are 246,683 Carmichael numbers below 10^{16}. Below 10^{16}, there are 279,238,341,033,925 primes; so there is less than a one-in-a-billion chance that a number is a Carmichael number.

We could make Fermat's Test into a perfect test, one that rejects every composite number, if there were some way to easily distinguish a prime from a Carmichael number. Unfortunately, to date nobody has developed a test that does not involve factoring the number. To develope one, the following propositions, given without proof, might be useful.

Proposition 3-4. *Every Carmichael Number is square free.*

Proposition 3-5. *If a prime p divides a Carmichael number n, then*

$$n \equiv 1 \pmod{p-1}.$$

4. Strong Pseudoprimality Test. We can improve Fermat's Test with an algorithm based on the following theorem given by G. Miller [**5**, p. 302].

Theorem 4-1. *Let n be an odd prime, and write n in the form $1 + 2^s d$ where d is odd. Then the Miller–Rabin sequence*

$$a^d, \ a^{2d}, \ a^{4d}, \ \ldots, \ a^{2^{s-1}d}, \ a^{2^s d} \pmod{n} \tag{4-1}$$

ends with 1; moreover, if a^d is not congruent to 1 (mod n), then the value directly preceding the first appearance of 1 is $n - 1$.

Proof: The assertion can be restated as follows: if n is prime and $n - 1 = 2^s d$, then Sequence 4-1 has the form,

$$(1, 1, \ldots, 1),$$

or the form,

$$(*, *, \ldots, n-1, 1, \ldots, 1)$$

for any a such that $1 < a < n$. By Fermat's Little Theorem, $a^{2^s d} \equiv 1 \pmod{n}$. Since n is prime, the only solutions to $x^2 \equiv 1 \pmod{n}$ are $x \equiv \pm 1 \pmod{n}$. Hence, it is easy to show that when $a^{2^y d} \equiv 1 \pmod{n}$, then $a^{2^{y-1} d}$ is congruent to either 1 or $n-1$. $\quad\square$

Theorem 4-1 also suggests a concept of pseudoprime.

Definition 4-2. If a composite n has the characteristics described in Theorem 4-1 for some base a, then n is called a *strong pseudoprime to the base a*. If n is either a prime or a pseudoprime, then n is called *probably prime*.

To implement Theorem 4-1, we reformulate it as follows.

Proposition 4-3 (Strong Pseudoprimality Test). *If $n - 1 = 2^s d$ with d odd and s nonnegative, then n is probably prime if $a^d \equiv 1 \pmod{n}$ or $(a^d)^{2^r} \equiv (n-1) \pmod{n}$ for some nonnegative r less than s.*

We now have a test that is even stronger than Fermat's Test; in fact, it reduces the number of pseudoprimes by half! Let $P(x)$ be the probability that a random odd integer x is composite and a base-a pseudoprime by Fermat's Test, and let $S(x)$ be the probability that x is a strong pseudoprime, then

$$S(x) = 4^{1-k} P(x)/(1 - P(x)).$$

Paul Erdös and Carl Pomerance [**4**, p. 278] proved that, as x goes to infinity, $S(x)$ goes to zero. So the larger x is, the better the Strong Test is. More important, there is no strong pseudoprime equivalent to the Carmichael number.

Consider the following proposition.

Proposition 4-4. *Let $n > 1$ be an odd composite integer. Then n passes the Strong Pseudoprimality Test for at most $(n-1)/4$ bases a with $1 < a < n$.*

Proof: We give only a sketch of this proof, as it is quite long. For a more detailed presentation, see [**7**, Sect. 8.4].

We assert that if p is an odd prime, then the number of incongruent solutions of

$$x^q \equiv 1 \pmod{p^\alpha}$$

is equal to the greatest common divisor of q and $p^{\alpha-1}(p-1)$. Let $n = 1 + 2^s d$ be a strong pseudoprime of base a. We then have

$$a^d \equiv 1 \pmod{n}$$

or

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

for some r less than s. In particular, we have $a^{n-1} \equiv 1 \pmod{n}$.

Let $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ be the prime factorization of n. We can now conclude that there are

$$\gcd(n - 1,\ p_i^{\alpha_i}(p_i - 1)) = \gcd(n - 1,\ p_i - 1)$$

incongruent solutions to

$$x^{n-1} \equiv 1 \pmod{p_i^{\alpha_i}}$$

for $i = 1, 2, \ldots, k$. Hence, by the Chinese remainder theorem [8, p. 67], there are only

$$\prod_{i=1}^{k} \gcd(n-1, \ p_i - 1)$$

incongruent solutions to

$$x^{n-1} \equiv 1 \pmod{n}.$$

Consider the case where at least one $p_j^{\alpha_j}$ has $\alpha_j \geq 2$. Then we have

$$\frac{p_j - 1}{p_j^{\alpha_j}} = \frac{1}{p_j^{\alpha_j - 1}} - \frac{1}{p_j^{\alpha_j}} \leq \frac{2}{9}$$

because p_j must be greater than or equal to 3. Then, for $n \geq 9$, we have

$$\begin{aligned}
\prod_{i=1}^{k} \gcd(n-1, p_i - 1) &\leq \prod_{i=1}^{k} (p_i - 1) \\
&\leq (\prod_{i \neq j}^{k} (p_i - 1))(\frac{2}{9} p_j^{\alpha_j}) \\
&\leq (\prod_{i \neq j}^{k} (p_i))(\frac{2}{9} p_j^{\alpha_j}) \\
&\leq (n-1)/4.
\end{aligned}$$

The case where n is square free can be found in [7]. Hence there are at most $(n-1)/4$ integers a less than $n-1$ when n is a strong pseudoprime.   □

If the Riemann hypothesis is true, then the number of bases a in which a composite number can be a strong pseudoprime is even more limited [5].

Conjecture 4-5. *If the generalized Riemann hypothesis is true, and if n passes the Strong Pseudoprimality Test for all integers a such that $1 < a < 2(\log n)^2$, then n is prime.*

We now have a theoretically perfect primality test. However, it is not reasonable to test n a total of $(n-1)/4$ times if n is very large. We run into a problem of efficiency like that with the Sieve of Eratosthenes. However, we can reasonably hope to test n for several bases. Let us examine the accuracy of such a test.

Proposition 4-6 (Miller–Rabin Probabilistic Primality Test). *Suppose we test an odd integer n for k randomly selected bases. If n is prime, then the result of the test is always correct. If n is composite, then the probability that n passes all k tests is at most $1/4^k$.*

Proof: The first assertion is obvious. For the second, consider the bases $a_1, \ldots, a_k$. By Theorem 4-4, the probability that a randomly selected base a allows a composite n to pass the Strong Test is at most $1/4$. Since the probability that the test fails for each a_i is at most $1/4$, the probability that an erroneous answer is chosen with k bases is at most $1/4^k$.   □

We can now see how truly powerful the Strong Test is. If we test only 100 bases, then the probability that we incorrectly identify a prime is less than 10^{-60}. It is more likely that the computer running the algorithm will produce errors, than that the Strong Test will fail for 100 tests [**10**, p. 53]. As we will shortly see, the Strong Test, even in its probabilistic form, may prove extremely useful for exactly proving primality efficiently.

5. Lucas Primality Test. The Lucas Primality Test is the most complicated test that we will describe in detail. Its basis is the notion of Lucas sequences, which are defined as follows.

Consider the quadratic equation $x^2 - ax + bx = 0$. Consider its descriminant $D = a^2 - 4b$, and its two roots, α and β. We have

$$\alpha = (a + \sqrt{D}D)/2 \text{ and } \beta = (a - \sqrt{D})/2.$$

We obtain the following three relations among a, b, D, α, and β:

$$\alpha + \beta = a, \ \alpha - \beta = D, \ \alpha\beta = b.$$

We can now define the Lucas sequences $U(a, b)$ and $V(a, b)$ using these relations when a, b, and D are nonzero integers.

Definition 5-1. For each $k \geq 0$, form the integers,

$$U_k(a, b) = \frac{\alpha^k - \beta^k}{\alpha - \beta} \text{ and } V_k(a, b) = \alpha^k + \beta^k.$$

The *Lucas sequences* of the pair (a, b) are the sequences,

$$U(a, b) = (U_0(a, b), U_1(a, b), U_2(a, b), \dots) \text{ and}$$
$$V(a, b) = (V_0(a, b), V_1(a, b), V_2(a, b), \dots).$$

An example of a Lucas sequence is the sequence of Fibonacci numbers; it is produced by $U(a, b)$ with $a = 1$ and $b = -1$. Like the Fibonacci numbers, Lucas sequences can be defined recursively: if $k \geq 2$, we have

$$U_k(a, b) = aU_{k-1} - bU_{k-2} \text{ and } V_k(a, b) = aV_{k-1} - bV_{k-2}.$$

We can now apply Lucas sequences to primality testing. The basic way is to replace the Strong Test's a^{n-1} with a Lucas sequence. Below, $(\frac{a}{p})$ is the Jacobi symbol, defined as follows. For p prime, set

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{if } x^2 \equiv a \pmod{p} \text{ is soluble;} \\ -1, & \text{if } x^2 \equiv a \pmod{p} \text{ is not soluble.} \end{cases}$$

For n composite with prime decomposition $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$, set

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{\alpha_1} \left(\frac{a}{p_2}\right)^{\alpha_2} \cdots \left(\frac{a}{p_k}\right)^{\alpha_k}.$$

We can now formally state the Lucas Theorem.

Theorem 5-2 (Lucas Theorem). *Let a, b, D, and U_k be as above. If p is an odd prime, if $\gcd(b, p) = 1$, and if $(\frac{D}{p}) = -1$, then p divides U_{p+1}.*

Proof: See [**6**, p. 104] for a detailed proof. $\square$

As before, for the actual primality test, we use the contrapositive of the Theorem.

Theorem 5-3 (Lucas Test). *Let n be an odd positive integer. If n does not divide U_{n+1}, then n is composite.*

As with the above tests, we have pseudoprimes for the Lucas Test.

Definition 5-4. If n is composite, if the greatest common divisor of n and b is equal to 1, if $(\frac{D}{n}) = -1$ and if n divides U_{n+1}, then n is called a *Lucas pseudoprime with parameters a and b.*

We now show how the Lucas Test can be modified to produce the Fermat Test. Let $\delta(n) = n - (\frac{D}{n})$. By Theorem 5-2, when n is prime and $\gcd(b, n) = 1$, we have $U_{\delta(n)} \equiv 0 \pmod{n}$.

Consider the case where $\frac{D}{n} = 1$. Let $D = \omega^2 \pmod{n}$. Then the roots modulo n are

$$\alpha = \frac{a + \omega}{2}, \text{ and } \beta = \frac{a - \omega}{2}.$$

Hence, the $(n-1)$th term of the corresponding Lucas sequence is

$$U_{n-1} = \frac{\alpha^{n-1} - \beta^{n-1}}{\alpha - \beta} = \frac{(\frac{a+\omega}{2})^{n-1} - (\frac{a-\omega}{2})^{n-1}}{w}.$$

Let $a = \omega + 2$. Then

$$U_{n-1} = \frac{(w + a)^{n-1} - 1}{w}.$$

Hence, $U_{n-1} \equiv 0 \pmod{n}$ implies

$$(w + 1)^{n-1} \equiv 1 \pmod{n}. \tag{5-1}$$

Equation 5-1 is the same as the Fermat Test. Hence, when referring to the Lucas Test, we will always assume $(\frac{D}{n}) = -1$, though the theorem is also true for $(\frac{D}{n}) = 1$.

We now will consider how to find a successful D for which $(\frac{D}{n}) = -1$. Ballie and Wagstaff have proposed two such methods [**1**].

The first method assigns D to the first element in the sequence $5, -7, 9, -11, \dots$ such that $(\frac{D}{n}) = -1$, $a = 1$, and $b = (1 - D)/r$. The second method assigns D to the first element in the sequence $5, 9, 13, \dots$ such that $(\frac{D}{n}) = -1$, a is the least odd greater than $D^{1/2}$, and $b = (a^2 - D)/4$. As each of these methods result in different parameters a and b, they produce different pseudoprimes. Both methods have their advantages.

For example, the first method is generally faster than the second. However, the second may prove to be far more important. Currently, there is no number known to be both a Carmicheal number and a Lucas pseudoprime. In other words, the combination of a Lucas Test and a Strong Pseudoprimality Test may always result in a prime. For the second method, it has been proved that the first 50 Carmichael numbers and several other Fermat pseudoprimes can never be Lucas pseudoprimes.

Conjecture 5-5. *Let n be a positive integer greater than 1. If n passes both a Strong Test and a Lucas Test, then n is prime.*

The above conjecture is used by Maple V Release 3 and later for its `isprime` function. This function can be used to find out if a number is indeed a prime. However, the test is a perfect test only if Conjecture 5-5 is true. Currently, we can make only the following claim to its exactness.

Proposition 5-6. *For any random positive integer n, Maple's* `isprime` *function will correctly determine if it is certainly prime with a probability of error* $\ll 10^{-15}$.

Proof: When testing a number for primality, the `isprime` function first performs the Strong Pseudoprimality Test up to a maximum of 25 times, randomly choosing a different basis a each time. If at any time n is found to be composite, the algorithm terminates. By Proposition 4-6, we know the probability that a composite n will pass 25 tests is less than or equal to $4^{-25} \leq 10^{-15}$. If n passes all 25 tests, Maple then performs a Lucas Test. We know that most strong pseudoprimes are not also Lucas pseudoprimes. Hence, the probability that `isprime` fails is much smaller than 10^{-15}.

$\square$

REFERENCES

[1] Ballie, R. J., and Wagstaff, Jr., S. S., "Lucas pseudoprimes," Mathematics of Computation (1980), 1391–417.

[2] Burton, D. M., "Elementary number theory," Allyn and Bacon, 1980.

[3] Dickson, L. E., "History of the theory of numbers," Vol. I, G. E. Stechert & Company, 1934.

[4] Erdös, P., and Pomerance, C., "On the number of false witnesses for a composite number," Math. Comp. (1986), 259–79.

[5] Miller, G., "Riemann's hypothesis and test for primality," Journal of Systems and Computer Science (1976), 300–17.

[6] Riben, T., "The new book of prime number records," 3rd ed., Springer-Verlag, 1995.

[7] Rosen, K., "Elementary number theory and its applications," 2nd ed., Addison-Wesley, 1993.

[8] Silverman, J. H., "A friendly introduction to number theory," Prentice-Hall, 1997.

[9] Simmons, G., "Calculus gems," McGraw Hill Inc., 1992.

[10] Tan, S. Y., "Perfect, amicable, and sociable numbers," World Scientific, 1996.

This page will be blank.