JOURNAL OF NUMBER THEORY 12, 128–138 (1980)

Probabilistic Algorithm for Testing Primality

MICHAEL O. RABIN

Institute of Mathematics, Hebrew University, Jerusalem, Israel, and Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Communicated by H. Zassenhaus

Received December 10, 1977

We present a practical probabilistic algorithm for testing large numbers of arbitrary form for primality. The algorithm has the feature that when it determines a number composite then the result is always true, but when it asserts that a number is prime there is a provably small probability of error. The algorithm was used to generate large numbers asserted to be primes of arbitrary and special forms, including very large numbers asserted to be twin primes. Theoretical foundations as well as details of implementation and experimental results are given.

The problem of determining for given integers whether they are prime has occupied mathematicians for centuries. In the words of Gauss in his celebrated "Disquisitiones Arithmeticae" [2, p. 396]:

> The problem of distinguishing prime numbers from composite numbers and of resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. It has engaged the industry and wisdom of ancient and modern geometers to such an extent that it would be superfluous to discuss the problem at length. Nevertheless we must confess that all methods that have been proposed thus far are either restricted to very special cases or are so laborious and prolix that even for numbers that do not exceed the limits of tables constructed by estimable men, i.e., for numbers that do not yield to artificial methods, they try the patience of even the practiced calculator. And these methods do not apply at all to larger numbers.

Gauss proceeds to give several interesting algorithms for the determination of primality. We do not intend to present here a survey or bibliography of the extensive literature concerning testing for primality. All the existing methods require for arbitrary integers n a number of steps $O(n^\epsilon)$ where ϵ is some fraction such as $\frac{1}{3}$ or $\frac{1}{4}$. Thus, experimental results indicate that around $n = 10^{65}$ each of these methods encounters some numbers untestable by it.

For integers of certain special forms, the most notable example being the so-called Mersenne numbers $n = 2^p - 1$, where p is a prime, there do exist

0022-314X/80/010128-11$02.00/0

rather rapid tests for primality. But arbitrary large integers defy even the most powerful computers, when tested by the conventional methods.

In this paper we present a primality test applicable to arbitrary and very large integers. If we count an arithmetical operation such as addition, multiplication, or division, involving integers $0 \leqslant a, b \leqslant n$ as one step then the test of primality of n requires in the worst case $c(\log_2 n)^2$ steps where c is about 100. Thus very large numbers are amenable to this test. In practice the number of steps in most cases is about $c \log_2 n$ for n which are prime, and about $3 \log_2 n$ for numbers n that are composite even if they have no small divisors.

The salient features of our method are that it is *probabilistic*, i.e., uses randomization within the computation, and that it produces the answer with a certain controllable miniscule probability of error. To be more precise, with full details to be given later, the algorithm produces and employs certain random integers $0 < b_1, \ldots, b_k < n$. If the outcome of the whole test is that n is composite then it is always correct (barring computational error). If the outcome asserts that n is prime then it may sometimes be wrong. But the probability that a composite number will be erroneously asserted to be prime is smaller than $1/2^{2k}$. If, say, $k = 50$, then the probability of error is at most $1/2^{100}$.

This last statement does *not* mean that an integer n asserted as prime by use of 50 random numbers is prime with probability at least $1 - 1/2^{100}$. Such an interpretation is nonsensical since n is either prime or not. The correct meaning is that if the test is applied to $m = 2^{100}$ integers $n_1, n_2, \ldots, n_m$, then the expected number of wrong answers is one. These integers need not be pairwise different and no probability distribution on the integers to be tested for primality is assumed.

Thus when we say that an integer n was asserted to be prime by use of 50 random numbers, this is no proof that n is actually prime. What is stated is that n was asserted to be prime by a procedure that on the average will make no more than one mistake in 2^{100} applications (even when testing the same n).

The test and an earlier, slightly weaker, form of the theorem underlying it appeared in [5]. A different probabilistic test for primality was given by Solovay and Strassen [6].

In Section 3 we discuss the implementation and application of the test to the generation of very large numbers asserted to be primes of arbitrary or of a desired prescribed form. Large primes are useful in exact arithmetical computations involving large integers. By computing modulo a prime, one can avoid the use of fractions. Also, recently evolved digitalized signature systems are based on the easy availability of large primes.

In the last section we present some of the experimental results which include examples of very large numbers asserted to be primes, twin primes, etc., very rapidly obtained on a medium-sized computer.

1. The Fundamental Theorem

Throughout this paper (a, b) denotes the greatest common divisor (g.c.d.) of the integers a, b, and $\mathrm{res}(a, b)$ denotes the least nonnegative residue of a when divided by b. By $b \mid a$ we denote the fact that b divides a, i.e., $\mathrm{res}(a, b) = 0$.

Definition. Let n be an integer. Consider the following condition to be denoted by $W_n(b)$, on an integer b:

(i) $1 \leqslant b < n$;

(ii)(a) $b^{n-1} \not\equiv 1 \bmod n$, or

(b) $\exists i$ s.t. $2^i \mid (n-1)$ and $1 < (b^{(n-1)/2^i} - 1, n) < n$.

Call such an integer b a *witness to the compositeness* of n.

This condition was first considered by Miller [4], who used it to give a nonprobabilistic test for primality assuming the correctness of the extended Riemann hypothesis. Our test does *not* assume *ERH* and is considerably faster.

Obviously if $W_n(b)$ holds for some b then n is composite. For if (ii)(a) holds then Fermat's relation is violated. And (ii)(b) means that n has a proper divisor. Thus the adjective "witness to compositeness" is justified. It turns out that if n is composite then witnesses abound.

Theorem 1. *If* $4 < n$ *is composite then*

$$\frac{3(n-1)}{4} \leqslant c(\{b \mid W_n(b)\}). \tag{1}$$

By $c(S)$ *we denote the number of elements in the set* S.

Because of (i), (1) means that no more than $\frac{1}{4}$ of the numbers $1 \leqslant b < n$ are *not* witnesses. We require some lemmas for the proof of Theorem 1.

Denote by E_n the set of all c, $1 \leqslant c < n$, $(c, n) = 1$. We have $c(E_n) = \phi(n)$, where ϕ is Euler's function. The set E_n is a group under multiplication mod n. If $n = p^k$, where p is prime, then $\phi(n) = p^k - p^{k-1}$; for odd p, E_n is a cyclic group.

For an integer m denote by $e(m)$ the largest i such that $2^i \mid m$.

For a sequence $m = \langle m_1, \ldots, m_k \rangle$ of integers define $\mathrm{res}(b, m) = \langle s_1, \ldots, s_k \rangle$, where $s_i = \mathrm{res}(b, m_i)$, $1 \leqslant i \leqslant k$. When m is fixed abbreviate $\mathrm{res}(b, m) = \mathrm{res}(b)$.

Lemma 2. *Let* $m_i \mid n$ *for* $1 \leqslant i \leqslant k$, *and* $(m_i, m_j) = 1$, $1 \leqslant i < j \leqslant k$, *and let* $m = \langle m_1, ..., m_k \rangle$. *For every* $s = \langle s_1, ..., s_k \rangle$, $s_1 \in E_{m_1}, ..., s_k \in E_{m_k}$, *the number of elements in*

$$E_n \cap \{b \mid \operatorname{res}(b, m) = s\}$$

is the same.

Proof. Let t be the product of all primes dividing n, but not any of the m_i, $1 \leqslant i \leqslant k$. Assume $t \neq 1$, for $t = 1$ the proof that follows has to be just slightly modified. By the Chinese Remainder Theorem there exists an integer b such that $\operatorname{res}(b, \langle m, t \rangle) = \langle s_1, ..., s_k, 1 \rangle$. Since $m_i \mid n$, $1 \leqslant i \leqslant k$, and $t \mid n$, we may assume $1 \leqslant b < n$. Now $b \in E_n$. Otherwise for some prime, $p \mid n$ and $p \mid b$. This p divides some m_i, or $p \mid t$. Since $s_i = b - q_i m_i$, if $p \mid m_i$ then $p \mid s_i$ contradicting $(m_i, s_i) = 1$. Similarly $p \mid t$ leads to the contradiction $p \mid 1$.

Denote the restriction $\operatorname{res}(\ , m) \mid E_n$ by f. By the previous paragraph, $f: E_n \to E_{m_1} \times \cdots \times E_{m_k} = G$ is a homomorphism of E_n *onto* the direct product G. Thus $f^{-1}(\langle s_1, ..., s_k \rangle)$ has the same number of elements for all $\langle s_1, ..., s_k \rangle \in G$.

Let, for example, $p \mid n$, $q \mid n$, where p and q are different primes. The number of $b \in E_n$ for which $\operatorname{res}(b, \langle p, q \rangle) = \langle s_1, s_2 \rangle$ is the same for every pair $1 \leqslant s_1 < p - 1$, $1 \leqslant s_2 < q - 1$. We shall employ in the sequel such considerations without further elaboration.

LEMMA 3. *Let* $p_1 \neq p_2$ *be primes*, $q_1 = p_1^{k_1}$, $q_2 = p_2^{k_2}$. *Assume* $q_1 q_2 \mid n$. *Denote* $t_i = (\phi(q_i), n - 1)$, $m_i = \phi(q_i)/t_i$, $i = 1, 2$.

At most $\phi(n)/m_1 m_2$ *of the integers* $b \in E_n$ *do not satisfy* $W_n(b)$.

If t_1 *is even then at most* $\phi(n)/2m_1 m_2$ *of the* $b \in E_n$ *do not satisfy* $W_n(b)$.

Proof. Let a_i be a primitive root mod q_i, $i = 1, 2$. That is, $a_i^t \equiv 1 \bmod q_i$ if and only if $\phi(q_i) \mid t$.

Let $b \in E_n$ then $(b, q_i) = 1$ and $b \equiv a_i^{r_i} \bmod q_i$, for some r_1, r_2. Because $q_1 q_2 \mid n$ we have that

$$b^{n-1} \equiv 1 \qquad \bmod n \tag{2}$$

implies $\phi(q_i) \mid r_i(n - 1)$, $i = 1, 2$. Hence it implies $m_i \mid r_i$, $i = 1, 2$, so $r_1 = h_1 m_1$, $r_2 = h_2 m_2$.

Thus if $\operatorname{res}(b, \langle q_1, q_2 \rangle) = \langle s_1, s_2 \rangle$ then for at most $1/m_1 m_2$ out of all pairs $\langle s_1, s_2 \rangle$ will (2) hold. By Lemma 2 all pairs $\langle s_1, s_2 \rangle$ appear equally often as residue of $b \in E_n$, hence (2) will hold for at most $\phi(n)/m_1 m_2$ of the integers $b \in E_n$. But if (2) does not hold then $W_n(b)$ is true.

The sharper claim made, when t_1 is even, hinges on the fact, to be proved next, that in this case even if (2) holds for b, in at least half the instances (ii)(b) holds, so that b is still a witness.

Assume that t_1 is even and $e(t_1) = e(t_2)$; see the notation preceeding Lemma 2. Let i be such that $(n - 1)/2^i = d(i)$ is an integer and $t_j \nmid d(i)$, $t_j/2 \mid d(i)$, $j = 1, 2$. Let $b \in E_n$ satisfy (2) and adopt the above notation, concerning a_1, a_2, r_1, r_2, etc. If h_1 is odd and h_2 is even then $\phi(q_1) \nmid h_1m_1d(i)$ and $\phi(q_2) \mid h_2m_2d(i)$. Hence

$$b^{d(i)} \not\equiv 1 \quad \bmod q_1, \tag{3}$$

$$b^{d(i)} \equiv 1 \quad \bmod q_2. \tag{4}$$

For such a $b \in E_n$ (ii)(b) holds. Similarly if h_1 is even and h_2 is odd. There are four equally frequent possibilities for the parities of h_1 and h_2. Thus for half of the $b \in E_n$ satisfying (2), $W_n(b)$ does still hold. Hence at most $\phi(n)/2m_1m_2$ integers $b \in E_n$ do not satisfy $W_n(b)$.

Finally let t_1 be even, $e(t_1) > e(t_2)$. Choose i and j so that $t_2 \mid d(i)$, $t_1 \nmid d(i)$, and $t_1/2^j \mid d(i)$ where j is minimal with the last property. Still using the above notation for a $b \in E_n$ satisfying (2), we have $\phi(q_2) \mid h_2m_2d(i)$ so that (4) holds. But $b^{d(i)} \equiv 1 \bmod q_1$ holds only if $\phi(q_1) \mid h_1m_1d(i)$, i.e., only if $2^j \mid h_1$. Unless $2^j \mid h_1$ we have (3), so that $W_n(b)$ is not true for at most $\phi(n)/2^jm_1m_2$ integers $b \in E_n$. But $1 \leqslant j$.

Proof of Theorem 1. Let $4 < n$ be composite. If $1 \leqslant b < n$ and $(b, n) \neq 1$ then (ii)(a) and consequently $W_n(b)$ holds. So that to prove (1) it suffices to show that at least $3\phi(n)/4$ of the $b \in E_n$ satisfy $W_n(b)$.

Assume that $n = p^k$, $1 < k$. Here $\phi(n) = p^{k-1}(p - 1)$ and $p \nmid (n - 1)$, so $p^{k-1} \leqslant \phi(p^k)/(\phi(p^k), n - 1) = m$. By the argument at the beginning of the proof of Lemma 3 applied to a single prime, we have that at most $\phi(n)/m$ of the $b \in E_n$ satisfy (2) and hence satisfy *not* $W_n(b)$. If $9 < n$ then $5 \leqslant p$ or $2 < k$ hence $4 \leqslant m$ and (1) follows. If $n = 9$, there are only two non-witnesses.

Next let n have at least two different prime divisors p_1, p_2, let $q_i = p_i^{k_i}$ be the maximal powers such that $q_1 \mid n$, $q_2 \mid n$, and assume that, say, $\phi(q_1) \nmid (n - 1)$. Then $2 \leqslant \phi(q_1)/(\phi(q_1), n - 1) = m_1$. If $\phi(q_2) \nmid (n - 1)$ then $2 \leqslant m_2$, and we have finished by the first statement of Lemma 3. Thus assume $\phi(q_2) \mid (n - 1)$. If $p_2 \neq 2$ then $t_2 = \phi(q_2)$ is even and by the second assertion of Lemma 3 at most $\phi(n)/2m_1 \leqslant \phi(n)/4$ (here $m_2 = 1$) of the $b \in E_n$ satisfy *not* $W_n(b)$. If $p_2 = 2$ then $\phi(n) \leqslant (n - 1)/2$ and at most $\phi(n)/m_1 \leqslant (n - 1)/4$ of the $1 \leqslant b < n$ satisfy (2).

Finally let $n = p_1^{k_1}p_2^{k_2} \cdots p_r^{k_r}$, $2 \leqslant r$, satisfy $\phi(p_i^{k_i}) \mid (n - 1)$, $1 \leqslant i \leqslant r$. Then $k_i = 1$, $1 \leqslant i \leqslant r$. An easy calculation shows $3 \leqslant r$. Such numbers do exist, e.g., (5), (6), and satisfy (2) for all $b \in E_n$. In honor of their discoverer [1] we call them Carmichael numbers. Let p_1, p_2, p_2 be three different prime divisors of n, $(p_i - 1) \mid (n - 1)$, $1 \leqslant i \leqslant 3$.

Assume $e(p_1 - 1) = e(p_2 - 1) = e(p_3 - 1)$, and let i be such that for $(n - 1)/2^i = d(i)$ we have $(p_j - 1) \nmid d(i)$, $(p_j - 1)/2 \mid d(i)$, $1 \leqslant j \leqslant 3$. If $b = a_i^{r_i} \bmod p_i$, where a_i is a primitive root mod p_i, $1 \leqslant i \leqslant 3$, then by the analysis in the proof of Lemma 3, $1 < (b^{d(i)} - 1, n) < n$ unless the r_i are all even or all odd. For if, say, r_1 is even and r_2 is odd then $p_1 \mid (b^{d(i)} - 1)$ and $p_2 \nmid (b^{d(i)} - 1)$. Thus *not* $W_n(b)$ holds for at most $\frac{2}{8}\phi(n)$ of the $b \in E_n$.

The similar analysis of the cases $e(p_3 - 1) = e(p_2 - 1) < e(p_1 - 1)$ and $e(p_3 - 1) < e(p_2 - 1) \leqslant e(p_1 - 1)$, is left to the reader.

If $n = p \cdot q \cdot r$ is a Carmichael number with $p \equiv q \equiv r \equiv -1 \bmod 4$, $p < q < r$. Then $n \equiv -1 \bmod 4$. Consequently $4 \nmid (n - 1)$ and i in (ii)(b) is just $i = 1$. Thus $W_n(b)$ holds for all $b \notin E_n$, i.e., for the

$$n\left(1 - \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right)\left(1 - \frac{1}{r}\right)\right) \leqslant n\left(1 - \left(1 - \frac{1}{p}\right)^3\right) \leqslant \frac{3n}{p}$$

numbers $(b, n) \neq 1$, and also for *exactly* $\frac{3}{4}\phi(n)$ of the $b \in E_n$.

Adding up:

$$c(\{b \mid W_n(b)) \leqslant \left(\frac{3}{4} + \frac{3}{p}\right)(n - 1).$$

If Carmichael numbers of the above type with arbitrarily large smallest prime factor do exist, then (1) is asymptotically best possible.

A systematic computer search for Carmichael numbers of this form was conducted by Oren by looking at triplets of primes $100 \leqslant p, g, r \leqslant 3000$. Many instances were found.

For example, for

$$652969351 = 271 \cdot 811 \cdot 2971 \tag{5}$$

the fraction of witnesses is 0.7513. For

$$2000436751 = 487 \cdot 1531 \cdot 2683 \tag{6}$$

the fraction of witnesses is 0.7507. Thus the constant $\frac{3}{4}$ in Theorem 1 seems to be best possible.

2. The Algorithm

Given a number n we choose a k, determined by the desired reliability, and randomly pick $1 \leqslant b_1, \ldots, b_k < n$.

Let $n - 1 = 2^l m$, where m is odd. For $b = b_i$ compute $b^m \bmod n$. This can be done by about $1.5 \log_2 m$ multiplications, see [2, p. 398]. Every time a number $n < d$ is obtained, find $d_1 = \text{res}(d, n)$ and continue. Thus only products of numbers $a, b < n$ and divisions of $d < n^2$ by n are involved. Next calculate the residues mod n of $b^{2m}, \ldots, b^{2^l m} = b^{n-1}$.

Since $\log_2 n = l + \log_2 m$, the evaluation of all the necessary powers of b will require $1.5 \log_2 n$ multiplications of a, $b < n$ and $1.5 \log_2 n$ divisions of a $d < n^2$ by n, altogether $3 \log_2 n$ steps.

If the residue of b^{n-1} is not 1 then $W_n(b)$ holds. If that residue is 1, find $(\text{res}(b^{2^i m}, n) - 1, n)$ for $1 \leqslant i \leqslant l$. Each g.c.d. computation can be performed by doing at most $\log_2 n$ subtractions and divisions by 2. If any one of these g.c.d.'s is neither 1 nor n then $W_n(b)$ is true. Thus for each b_i we computationally determined whether $W_n(b_i)$ is true, and the total required number of steps is $3 \log_2 n + l \cdot \log_2 n$.

If, for any $1 \leqslant i \leqslant k$, $W_n(b_i)$ is true then n is composite. If, for all $1 \leqslant i \leqslant k$, $W_n(b_i)$ is not true, then the test asserts that n is prime.

Theorem 2. *The above algorithm requires for $n - 1 = 2^l m$, m odd, at most $k \cdot (2 \cdot \log_2 n + l \cdot \log_2 n)$ steps. If n is prime then the result is always correct. For each composite n the test may declare n to be prime, but the probability of such error is smaller than $1/2^{2k}$.*

Proof. Only the last statement requires proof. An error will occur only when the n to be tested is composite and the $b_1, \ldots, b_k$ chosen in this particular run of the algorithm are all nonwitnesses. Becauses of Theorem 1, the probability of randomly picking a nonwitness is smaller than $\frac{1}{4}$. The probability of independently picking k nonwitnesses is smaller than $1/4^k = 1/2^{2k}$. Thus the probability that for any given number n, a particular run of the algorithm will produce an erroneous answer is smaller than $1/2^{2k}$.

Note that the theorem is formulated with respect to each n, no averaging over a range of n's is necessary. For a given composite n the test may (rarely) give a wrong answer in one run and the correct answer in other runs. Thus there are no n's on which the algorithm behaves poorly.

3. Implementation

In implementing the algorithm we incorporate some laborsaving steps. First we test the given n for divisibility by any prime $p < N$, where, say $N = 1000$.

If n passed this sieve then we apply the algorithm. Suppose that we pick $k = 30$. An examination of the proof of Theorem 1 shows that for "most" composite n the probability of finding a witness in one try is much bigger than $\frac{3}{4}$. In fact, in actual runs when n was composite the witness was always the first or second b chosen. But even for n for which (1) is almost exact, the probability of error is $1/2^{60} < 10^{-18}$. One expected error in a billion billion tests! This seems small when compared to the frequency of machine errors present in practical computations.

The main application of the primality test is to produce large numbers asserted to be primes with specified properties.

The overall strategy is to generate numbers of the specified form and test for primality until a number is asserted to be prime. Such searches lead to success rather quickly because in most cases the Prime Number Theorem ensures that the density of primes in the sequence m_i searched is like $1/l$ if the numbers satisfy $\ln(m_i) \sim l$. In other cases an appropriate density statement can be derived from stronger hitherto unproven assumptions such as the extended Riemann hypothesis (ERH). In every instance actually tried, the search did terminate within practical time.

As mentioned before, any primality test should incorporate trying small divisors. Thus in practice the full test is applied only to the numbers asserted to be primes.

If desired, it is often possible to ensure that $n - 1 = 2 \cdot m$, where m is odd so that the second part of checking W_n, that relating to (ii)(b), involves only $b^{(n-1)/2}$. But even if this feature is not incorporated, a heuristic argument shows that the expected number of g.c.d. computations arising from (ii)(b) is two. Note that the notion of expected number used here does not have a precise meaning as in Theorem 2 and is invoked only heuristically.

Suppose we want to find an "arbitrary" prime having 250 binary digits. Start by randomly generating a sequence $m = \alpha_1\alpha_2 \cdots \alpha_{250}$, $\alpha_i \in \{0, 1\}$, $\alpha_1 = 1$. View m as an integer written in binary notation, say it is even. Successively form the sums $m + 1, m + 3, \ldots$; if desired omit those for which $4 \mid (n - 1)$. For each $n = m + i$ test for primality. Continue until a number is asserted to be prime.

Because of the Prime Number Theorem, the search will not require testing a prohibitive number of sums $m + i$. The density of primes around $n = 2^{250}$ is $1/\ln n > 1/250$. Thus, one can incorporate a stopping rule: If within 500 tries no prime was found, drop the number m and restart with a new number m, $[\log_2 m] = 250$. In this way a number asserted to be prime $n = m + i$, $i < 1000$, will usually be found very rapidly.

For certain applications it is important to have a prime n so that $n - 1$ has a large prime factor. This is achieved in a similar fashion. Suppose that we want $n - 1$ to have a prime factor p so that $[\log_2 p] = 200$. Find, as before, such a number asserted to be prime. Form the numbers $2ip + 1$, $i = 1, 3, \ldots$. Test each number for primality. Again a number $q = lp + 1$ of the desired form is rapidly asserted as prime. If the arithmetical series with difference p does not rapidly find a number asserted to be prime (this has never happened), then p can be discarded and another prime p_1 can be used for restarting the search.

These are just examples of searches actually conducted. The variations pertaining to other forms of primes, or to the testing of given numbers are obvious.

4. Some Experimental Results

V. Pratt has programmed this algorithm and together we planned some experiments. The computations were done on a medium-sized computer. Numbers with several hundred binary digits were generated and tested. Still, the computations, including searches, did not take more than minutes per number asserted to be prime.

The first test was verification of the known results concerning primality and compositeness of $n = 2^p - 1$, $p \leqslant 500$, p prime. Within about 10 min all answers were produced without a single error. This was mainly done to test the program. For Mersenne primes our test is about $1/k$ as fast as the Lucas test which applies only to Mersenne numbers.

Next the test was applied to generate some large numbers asserted to be primes along the lines explained in Section 3. Rather than load an arbitrary binary sequence m, the search started from powers of 2 and proceeded by decrements. The numbers

$$2^{300} - 153, \qquad 2^{400} - 593,$$

were asserted to be the largest primes below 2^{300} and 2^{400}, respectively, since the other numbers in the intervals are all composite.

Finally the test was applied to discover what we believe are the hitherto largest known numbers asserted to be twin primes. In order to speed up the sieving process the search was started with a number m which is a product of many small primes. Then pairs of the form $m \cdot l + 821$, $m \cdot l + 823$, and $m \cdot l + 827$, $m \cdot l + 829$ were tried. The pairs 821, 823 and 827, 829 are themselves twin primes. Within half an hour

$$\Big(\prod_{p_i<300} p_i\Big) \cdot 338 + 821, \qquad \Big(\prod_{p_i<300} p_i\Big) \cdot 338 + 823,$$

were asserted to be twin-primes. These numbers are of order of magnitude 10^{123}. Five subsequent hours of search failed to discover additional pairs. This is the only case where any search required more than a few minutes. Of course, no Prime Number Theorem density estimates apply to twin primes. The reader's guess as to the heuristic implications of this seeming gap in twin primes is as good as ours.

In conclusion, let us raise a question concerning possible theoretical applications of this primality test, in addition to the practical applications mentioned in the Introduction. What conjectures pertaining to the distribution of primes can one formulate, lend support to, or disprove by use of this test? For example, are there some refined estimates for the density of primes around 10^{100}, perhaps consequences of some strong number- or function-theoretic conjectures, which one could experimentally check using the primality test. The author, together with V. Pratt, have tested the density of primes of the form $n = x^2 + 1$ for $n \sim 2^{50}, 2^{100}, 2^{150}, 2^{200}$, and found very good agreement with the Hardy–Littlewood conjecture on this density. These and other experimental results will be reported elsewhere.

Note added May 10, 1978: Donald Knuth (and, it seems, several others) observed that the primality test can dispense with the g.c.d. computation. His observation is based on the following corollary of Theorem 1.

Using the notation of Section 2, let $n - 1 = 2^l \cdot m$, where m is odd. Let $0 \leqslant x < n$, denote $x_0 \equiv x^m \bmod n$, $x_i \equiv x_{i-1}^2 \bmod n$, $1 \leqslant i \leqslant l$. Thus $x_l \equiv x^{n-1} \bmod n$.

PROPOSITION. *If $4 < n$ is composite and odd then for at least $\frac{3}{4}(n - 1)$ of the $1 \leqslant x < n$ either $x_l \neq 1$ or for some $1 \leqslant i \leqslant l$ we have $x_i = 1$ and $x_{i-1} \neq n - 1$.*

Proof. By Theorem 1, if n is composite then for at least $\frac{3}{4}(n - 1)$ of the $1 \leqslant x < n$ $W_n(x)$ holds. If $W_n(x)$ holds then $x_l \neq 1$, or $x_l = 1$ and for some $0 \leqslant j \leqslant l$ we have $1 < (x_j - 1, n) < n$.

In the second case let $i - 1$, $j \leqslant i - 1 < l$, be the last index such that $x_{i-1} \neq 1$; thus $x_i = 1$. We have $x_j \equiv 1 \bmod p$, where p is a proper divisor of n. Hence $x_{i-1} \equiv 1 \bmod p$, since

$$x_{i-1} \equiv x_j^{2^{i-j-1}} \qquad \bmod n.$$

Now $x_{i-1} = n - 1$ is impossible because it entails $x_{i-1} - 1 = n - 2$ being divisible by p, but $p \neq 2$. Thus $x_{i-1} \neq n - 1$ and $x_i = 1$.

Note that if the condition in the statement of the proposition holds for some x then n must be composite. Namely, if $x_l \neq 1$ then Fermat's relation does not hold. And if $x_{i-1} \neq n - 1$, $x_i \equiv x_{i-1}^2 \equiv 1 \bmod n$, then 1 has more than two square roots mod n so n is composite.

The test for primality runs as in Section 2 except that when $x_l = 1$ we need not compute the $(x_i - 1, n)$. Simply calculate $x_0 \equiv x^m \bmod n$, then square mod n repeatelly until $x_{i-1} \neq n - 1$ and $x_i = 1$, or until x_l is reached. In any case we need never square x_{l-1}.

REFERENCES

1. R. D. CARMICHAEL, On composite numbers p which satisfy the Fermat congruence $a^{p-1} \equiv 1 \bmod p$, *Amer. Math. Monthly* **19** (1912), 22–27.
2. C. F. GAUSS, "Disquisitiones Arithmeticase" (A. A. Clarke, Transl.), Yale Univ. Press, New Haven, Conn. London, 1966.
3. D. E. KNUTH, "The Art of Computing," Vol. 2, Addison-Wesley, Reading, Mass., 1969.
4. G. L. MILLER, Reimann's hypothesis and a test for primality, *J. Comput. and System Sci.* **13** (1976), 300–317.
5. M. O. RABIN, Probabilistic algorithms, *in* "Algorithms and Complexity, Recent Results and New Direction" (J. F. Traub, Ed.), pp. 21–40, Academic Press, New York, 1976.
6. R. SOLOVAY AND V. STRASSEN, A fast Monte-Carlo test for primality, *SIAM J. Comput.* **6** (1977), 84–85.