Primality Tests

Most of our more-sophisticated primality checking will rely on two results—the Rabin-Miller test and the Pocklington test. The former is used as a “gatekeeper,” since it is only capable of proving compositeness, while the second is used as a final check, since it is more difficult to verify.

**Theorem 1 (Rabin-Miller):** If \( p = t2^s + 1 \) is prime, where \( t \) is odd (i.e., \( s \) is maximal), and \( 2 \leq a \leq p - 1 \), then either \( a^t \equiv 1 \pmod{p} \) or \( a^{t2^r} \equiv -1 \pmod{p} \) for some \( 0 \leq r \leq s - 1 \).

**Proof:** If \( p \) is prime and \( 2 \leq a \leq p - 1 \), then Fermat’s Little Theorem implies that \( a^{p-1} \equiv 1 \pmod{p} \). Furthermore, if \( p \) is prime, then the ring of integers modulo \( p \) is a field, which implies that 1 has at most two square roots, namely 1 and \(-1\) (which are equal if \( p = 2 \)). Consider, then, the sequence (modulo \( p \))

\[
a^{t2^s}, a^{t2^{s-1}}, a^{t2^{s-2}}, \ldots, a^t
\]

Each term in this sequence is one of the square roots of the preceding term, and the first term is 1. There are, then, two possibilities: either the whole sequence consists of 1’s, or some element of the sequence after the first is a \(-1\). These are precisely the two possibilities in the conclusion of the theorem. **Q.E.D.**

We also note (proof later) that a composite \( p \) will pass this test for at most \( 1/4 \) of the possible \( a \) values—a composite \( p \) that passes this test for the base \( a \) is said to be a *strong pseudoprime to the base* \( a \). We also note that performing a Rabin-Miller test for a particular \( a \) requires only a single modular exponentiation, followed by repeated squaring (up to \( s - 1 \) times). It is also true that calculation of \( s \) and \( t \) is trivial on a binary computer, since \( s \) is simply the number of trailing 0 bits in \( p - 1 \) and \( t \) is the result of shifting \( p - 1 \) right by \( s \) bits.

**Theorem 2 (Pocklington):** If \( p = q^kr + 1 \), \( q \) is prime, \( q \nmid r \), and there exists \( 2 \leq a \leq p - 1 \) such that \( a^{p-1} \equiv 1 \pmod{p} \) and \((a^{(p-1)/q} - 1, p) = 1\), then every prime factor of \( p \) is congruent to 1 modulo \( q^k \).

**Proof:** Let \( s \) be a prime factor of \( p \) and let \( m \) be the multiplicative order of \( a \) modulo \( s \) (i.e., the smallest positive integer such that \( a^m \equiv 1 \pmod{s} \)). The first condition on \( a \) ensures that \( m|p-1 = q^k r \), while the second ensures that \( m \nmid (p-1)/q = q^{k-1}r \). These two imply that \( q^k|m \). Now, Fermat’s Little Theorem implies that \( m|s-1 \), hence \( q^k|s-1 \), or \( s \equiv 1 \pmod{q^k} \). **Q.E.D.**

A couple of useful corollaries -

**Corollary 3:** Let \( p, q, r, a \) be as in **Theorem 2**. If, in addition, \( q^k > r \) then \( p \) is prime.

**Proof:** Using **Theorem 2**, we see that all prime factors of \( p \) are greater than \( q^k > \sqrt{p} \). Hence, \( p \) is prime. **Q.E.D.**
Corollary 4: Let \( p, q, r, a \) be as in Theorem 2. If, in addition, \( q^{2k} > r \) then either \( p \) is prime or is the product of two primes congruent to 1 modulo \( q^k \).

Proof: Using Theorem 2, we see that all prime factors of \( p \) are greater than \( q^k > \sqrt[3]{p} \). Hence, there are at most 2 of them and they are both congruent to 1 modulo \( q^k \). Q.E.D.

Corollary 5: Let \( p, q, r, a \) be as in Corollary 4. Assume that \( p \) and \( q \) are both odd.

Let \( r = bq^k + c \) where \( 0 \leq c < q^k \). If \( b \) is not a multiple of 4 or \( c^2 - 4b \) is not a square, then \( p \) is prime.

Proof: We only need to rule out the case where \( p = p_1p_2, p_i \) prime, \( p_i = k_iq^k + 1 \). First, observe that, since \( p_1p_2 = p \), we have \( k_1k_2 < q^k \). Furthermore, each \( k_i \) must be even and nonzero. Hence, we have \( 2 \leq k_i \leq (q^k - 1)/2 < q^k \). Furthermore, \( k_1 + k_2 \leq 2 + (q^k - 1)/2 < q^k \), since the sum of two real numbers of constant product is a maximum when one is as small as possible (the special case when \( q = 3, k = 1 \) is easily dealt with, since no two even positive integers have product less than 3). Thus, we see that

\[
(k_1q^k + 1)(k_2q^k + 1) = k_1k_2q^{2k} + (k_1 + k_2)q^k + 1
= bq^{2k} + cq^k + 1
\]

implies that \( k_1k_2 = b, k_1 + k_2 = c \). Since the \( k_i \) are both even, \( b \) must be a multiple of 4. Furthermore, \( c^2 - 4b = (k_1 - k_2)^2 \), so \( c^2 - 4b \) must be a square. Since one or the other of these was assumed to be false, the other conclusion of Corollary 4 must hold, namely, \( p \) must be prime. Q.E.D.

To prove that a composite \( p \) is a strong pseudoprime to at most 25% of the possible bases, we need two lemmas:

Lemma 6: In a cyclic group of order \( n \), there are \((n, k)\) distinct elements \( x \) that satisfy \( x^k = 1 \).

Proof: Let \( d = (n, k) \) and let the cyclic group be generated by \( g \), so that the group is \( \{g, g^2, g^3, \ldots, g^n = 1\} \). An element \( g^j \) satisfies the equation iff \( n|(jk) \) iff \( (n/d)|(jk/d) \) iff \( j \) is a multiple of \( n/d \) since \( n/d \) and \( k/d \) are relatively prime. There are \( d \) such values \( 1 \leq j \leq n \). Q.E.D.

Lemma 7: Let \( p = t^{2^s} + 1 \) be prime with \( t \) odd. Then, the number of \( 1 \leq x \leq p - 1 \) that satisfy \( x^{2^r} \equiv -1 \pmod{p} \) is 0 if \( r \geq s \) and \( 2^r(u, t) \) otherwise.

Proof: Let \( g \) be a generator for the multiplicative group of nonzero elements modulo \( p \) and let \( x = g^j \). Then, the number of distinct \( x \) that satisfy the condition is the same as the number of distinct exponents \( j \) that satisfy

\[
ju2^s \equiv (p - 1)/2 \pmod{p - 1} \equiv t^{2^s - 1} \pmod{t2^s}
\]

Clearly, if \( r \geq s \), this cannot occur since the left-hand side and the modulus both contain at least \( s \) factors of 2, while the right-hand side only has \( s - 1 \). On the other hand,
if \( r < s \), denote \((u, t)\) by \(d\). In this case, there is at least one solution since \((u/d)\) is relatively prime to \((t/d)^{2^{s-r}}\). This implies that there is a \(1 \leq k < (t/d)^{2^{s-r}}\) which is the multiplicative inverse of \((u/d)\) modulo \((t/d)^{2^{s-r}}\). Now, let \(j = k(t/d)^{2^{s-r}-1}\). Observe that \(j(u/d) \equiv (t/d)^{2^{s-r}-1}\) (mod \((t/d)^{2^{s-r}}\)) which implies that

\[
j u^2r \equiv t 2^{s-1} \pmod{t 2^s}
\]

Once we have one solution, we can easily count the others using Lemma 6, since all solutions will be a product of the one fixed solution and a solution of \(y^{u2^r} \equiv 1 \pmod{p}\). Thus, the total number of solutions is \((t 2^s, u^2r) = 2^r(u, t)\). Q.E.D.

**Theorem 8:** If \(p\) is odd and composite, it is a strong pseudoprime to at most \((p-1)/4\) bases \(0 < a < n\).

**Proof:** We will break this up into 3 cases –

**Case I:** \(p\) is divisible by the square of an odd prime \(q\). Suppose \(p\) is a strong pseudoprime relative to \(0 < a < p\), and \(q^k | p \) (\(k\) maximal), \(k \geq 2\). Then, \(a^{p-1} \equiv 1 \pmod{q^k}\). The size of the group in question, the multiplicative group of the integers modulo \(q^k\) is \(\varphi(q^k) = q^{k-1}(q-1)\). This tells us that, among the \(a\) less than \(q^k\), there are \(d = (q^{k-1}(q-1), p-1)\) solutions. Now, \(q\) is prime and \(q | p \) so \(q \nmid p-1\). Therefore, \(d | q-1\). Using the Chinese Remainder Theorem, then, we see that the number of such \(a\) is at most \((q-1)p/q^k\) and thus the proportion of solutions is at most

\[
\frac{(q-1)p}{q^k(p-1)} \leq \frac{(q-1)p}{q^k(p-(p/q^k))} \\
\leq \frac{(q-1)p}{pq^k-p} \\
= \frac{q-1}{q^k-1} \\
\leq \frac{q-1}{q^2-1} \\
= \frac{1}{1+q} \leq 1/4
\]

Note that this case does not really use the full strength of the Rabin-Miller test, only the Fermat portion.

**Case II:** \(p\) is the product of two distinct odd primes, \(p = q_1q_2\). Let \(q_1 = t_1 2^{s_1} + 1\) and \(q_2 = t_2 2^{s_2} + 1\) \((t_i \text{ odd})\). Suppose, without loss of generality, that \(s_1 \leq s_2\). Note that \(s_1 \leq s\) since

\[
t 2^s = p - 1 = (q_1 - 1)(q_2 - 1) + (q_1 - 1) + (q_2 - 1) = t_1 2^{s_1} t_2 2^{s_2} + t_1 2^{s_1} + t_2 2^{s_2} = 2^{s_1} (t_1 t_2 s_2 + t_1 + t_2 2^{s_2-s_1})
\]

The Chinese Remainder Theorem then lets us reinterpret the strong pseudoprime condition: if \(p\) is a strong pseudoprime to base \(a\), then either \(a^t \equiv 1 \pmod{q_1}\) and \(a^t \equiv 1 \pmod{q_2}\)
(mod $q_2$) or, for some $0 \leq r < s$, $a^{t_2^r} \equiv -1 \pmod{q_1}$ and $a^{t_2^r} \equiv -1 \pmod{q_2}$. Using Lemma 6, we see that the first condition holds for

\[(t, q_1 - 1)(t, q_2 - 1) = (t, t_1 2^{s_1})(t, t_2 2^{s_2})
= (t, t_1)(t, t_2)
\leq t_2 t_2\]

Next, Lemma 7 implies that, for $0 \leq r < s_1 \leq s_2$, that the second condition has

\[2^r (t, t_1)2^r (t, t_2) \leq 4^r t_1 t_2\]
solutions (there are none if $r \geq s_1$).

Thus, the total number of solutions is at most

\[t_1 t_2 (2 + 4 + 4^2 + \cdots + 4^{s_1-1})\]

Furthermore, $p - 1 > (q_1 - 1)(q_2 - 1) = t_1 t_2 2^{s_1 + s_2}$ so the proportion of solutions is at most

\[\frac{1 + \frac{4^{s_1-1}}{4-1}}{2^{s_1 + s_2}}\]

If $s_1 < s_2$, then this is at most

\[2^{-2s_1-1} \left(\frac{2}{3} + \frac{4^{s_1}}{3}\right) \leq 2^{-3} \frac{2}{3} + \frac{1}{6} = \frac{1}{4}\]

If $s_1 = s_2$, then we must be a bit more careful. We claim that, in this subcase, at least one of the $t_i$ is not a factor of $t$. For, if $t_1 | t$, then

\[p - 1 = t2^s
= q_1 q_2 - 1
= (q_1 - 1)q_2 + (q_2 - 1)
= t_1 2^{s_1} q_2 + t_2 2^{s_2}
= 2^{s_1} (t_1 q_2 + t_2)\]

so that $0 \equiv t_2 2^{s_1} \pmod{t_1}$, i.e. $t_1 | t_2$. Similarly, if $t_2 | t$, then $t_2 | t_1$. Thus, if both $t_i$ are factors of $t$, then they are equal and hence $q_1 = q_2$, a contradiction. So, at least one of the $(t_i, t)$ is strictly less than $t_i$, hence less than $t$ by at least a factor of 3. Recall that, in our counting of solutions, we replaced $(t_1, t)(t_2, t)$ by $t_1 t_2$. This argument shows that this was overly generous by at least a factor of 3, so we may now replace $t_1 t_2$ by $t_1 t_2/3$. This gives us the upper bound on the proportion of solutions of

\[2^{-2s_1} \left(\frac{2}{3} + \frac{4^{s_1}}{3}\right) \leq \frac{1}{18} + \frac{1}{9} = \frac{1}{6} < \frac{1}{4}\]
Case III: $p$ is the product of three or more distinct primes, $p = q_1 q_2 \ldots q_n$ ($n \geq 3$). Proceed as in Case II and let $q_i = t_i 2^{s_i} + 1$ with $t_i$ odd. Assume, without loss of generality that $s_i \leq s_{i+1}$. Arguing as before, we see that the proportion of solutions is at most

$$2^{-s_1 - s_2 - \cdots - s_n} \left( 1 + \frac{2^{ns_1} - 1}{2^n - 1} \right) \leq 2^{-ns_1} \left( \frac{2^n - 2}{2^n - 1} + \frac{2^{ns_1}}{2^n - 1} \right)$$

$$= 2^{-ns_1} \frac{2^n - 2}{2^n - 1} + \frac{1}{2^n - 1}$$

$$\leq 2^{-n} \frac{2^n - 2}{2^n - 1} + \frac{1}{2^n - 1}$$

$$= \frac{2 - 2^{1-n}}{2^n - 1}$$

$$= 2^{1-n}$$

$$\leq \frac{1}{4}$$

since $n \geq 3$.

Q.E.D.
Mihailescu’s Prime-Generation Algorithm

To generate a provable prime $p$ of $n$ bits, Mihailescu has (more or less) proposed the following algorithm which combines a number of the above results:

Step 0: if $n \leq 16$, return an appropriately-size prime from a list of the 16-bit primes.

Step 1: Recursively generate a prime $q$ of size at least $\lceil n/3 \rceil$.

Step 2: Set up a sieve with a start value of at least $\lceil (2^n - 1)/(2q) \rceil$ and a size of at least $10n$.

Step 3: For all 16-bit primes $r$, remove from the sieve all values $t$ such that $r|2qt + 1$. Note that this necessitates calculating $(2q)^{-1} \mod r$.

Step 4: If the sieve is empty, go back to Step 2 (set up a nonoverlapping sieve). Otherwise, for each sieve output $t$, perform a base-2 Rabin-Miller test on $p = 2qt + 1$. If it fails, go back to Step 4. If it passes, go on to Step 5.

Step 5: Divide $2t$ by $q$, and call the quotient $b$ and the remainder $c$. If $b$ is a multiple of 4, and $c^2 - 4b$ is a square, go back to Step 4. (For somewhat subtle number-theoretic reasons, it’s really only necessary to check whether or not $c^2 - 4b$ is a square – if it is, $b$ is necessarily a multiple of 4).

Step 6: Let $a$ denote a small prime (start with 2, continue to $L$). If you have reached $L$, go back to Step 4. Let $d$ denote $a^{2t} \mod p$. If $d = 1$, go back to Step 6 (next small prime). Otherwise, calculate $d^q \mod p$. If this is not 1, go back to Step 4. Calculate $(d - 1, p)$. If this is not 1, go back to Step 4. If it is 1, then $p$ is prime. Return it, and terminate the algorithm.

Note that Step 1 and Step 2 may be randomized so that different primes are produced each time.

A prime certificate is a list containing all information necessary for a third party to verify the calculations to prove primality. In this case, a certificate for $p$ would be:

1) $p$ itself,
2) $q$,
3) a prime certificate for $q$,
4) if $b$ is not a multiple of 4, then $b$, else $c^2 - 4b$ (to verify it’s not a square),
5) the $a$ value that finally worked in Step 6.