3.5 RABIN's Probabilistic Primality Test

RABIN transferred an idea of SOLOVAY and STRASSEN to MILLER's test. As it later turned out SELFRIDGE had used the method already in 1974.

If we choose a random base a in $[2 \dots n-1]$, then n "in general" fails the strong pseudoprime test to base a except when it is prime. But what means "in general"? How large is the probability? To answer this question we look at the corollary of Proposition 10 where the tighter bound $\frac{1}{4}$ was stated without proof.

Note that the bound $\frac{1}{4}$ is sharp. To see this we consider integers of the form

$$n = (1+2t)(1+4t)$$

with odd t (and assume that $p = 1+2t$ and $q = 1+4t$ are prime—example: $t = 24969$, $p = 49939$, $q = 99877$). Then $n-1 = 2r$ with $r = 3t+4t^2$, and

$$B_u = \{a \mid a^r \equiv 1 \pmod{n}\} \cup \{a \mid a^r \equiv -1 \pmod{n}\}.$$

Since $\gcd(r, p-1) = \gcd(3t+4t^2, 2t) = t = \gcd(r, q-1)$, each of these two congruences has exactly t^2 solutions. Hence $\#B_u = 2t^2$,

$$\frac{\#B_u}{n-1} = \frac{2t^2}{2 \cdot (3t+4t^2)} = \frac{t}{3+4t} = \frac{1}{4+\frac{3}{t}}.$$

However most composite integers don't even come close to this bound $\frac{1}{4}$.

In general assume we are given a family $(M_{(n)})_{n \geq 1}$ of sets $M_{(n)} \subseteq [1 \dots n-1]$ and a real number $\varepsilon \in]0,1[$ with

1. $M_{(n)} = [1 \dots n-1]$ if n is prime,
2. $\#M_{(n)} \leq \varepsilon \cdot (n-1)$ for all sufficiently large odd composite integers n.

Moreover we assume that the property $a \in M_{(n)}$ is efficiently decideable for all $a \in [1 \dots n-1]$, i. e. with costs that grow at most polynomially with $\log(n)$. Then we have a corresponding (abstract) pseudoprime test:

1. Choose a random $a \in [1 \dots n-1]$.
2. Check whether $a \in M_{(n)}$.
3. Output:
   (a) If **no**: n is composite.
   (b) If **yes**: n is pseudoprime to a.

The corresponding **probabilistic primality test** consists of a series of k of these pseudoprime tests to independently chosen bases a (note that this allows for accidental repetitions). If $a \notin M_{(n)}$, we call a a witness for

compositeness of n. If always $a \in M_{(n)}$ (we find no witnesses), then n is almost certainly a prime. We may assign an "error probability" δ to this event. This is computed in the following way (no it is $not = \varepsilon^k$):

Consider the set of odd r-bit integers, that is odd positive integers $< 2^r$. Let X be the subset of *composite* numbers, and Y_k, the subset of integers that pass the first k of a given series of independent (abstract) pseudoprime tests. The probability that a composite integer makes it into this subset is the conditional probability $P(Y_k|X) \leq \varepsilon^k$.

Nevertheless more important for the practical application is the "converse" probability $\delta = P(X|Y_k)$ that a number n that passed all the tests is still composite. This probability is assessed using BAYES' formula:

$$P(X|Y_k) = \frac{P(X) \cdot P(Y_k|X)}{P(Y_k)} \leq \frac{P(Y_k|X)}{P(Y_k)} \leq \frac{1}{q} \cdot \varepsilon^k \leq r \cdot \ln(2) \cdot \varepsilon^k,$$

where we also used the density of primes estimated by the prime number theorem:

$$P(Y_k) \geq P(\text{prime}) =: q \geq \frac{1}{r \cdot \ln(2)}$$

(the latter inequality being rather tolerant since we consider only odd numbers). Thus the "error probability" $\delta = P(X|Y_k)$ might be larger than ε^k. We can (and should) reduce it by restricting the set we search for primes, thereby enlarging $P(Y_k)$. For example before starting the series of pesudoprime tests we could try to divide by all primes say $< 100r$.

For RABIN's **primality test** we take $M_{(n)}$ as the set of bases n is a strong pseudoprime to, and $\varepsilon = \frac{1}{4}$. If n passes 25 single tests then it is prime with a quite small error probability. The probability that an exact computation produces a false result due to a hardware nor software error is larger than the error probability of RABIN's algorithm. KNUTH even doubts whether a future published proof of the extended RIEMANN hypothesis might ever be as trustworthy. Nevertheless from a mathematical viewpoint we are unsatisfied when we can't be sure that we really found a prime.

For further information on the error probability of a probabilistic primality test read

- S. H. KIM/C. POMERANCE: The probability that a random probable prime is composite. Math Comp. 53 (1989), 721–741.
- ALFRED J. MENEZES, PAUL C. VAN OORSCHOT, SCOTT A. VANSTONE: *Handbook of Applied Cryptography.* CRC Press, Boca Raton 1997, p. 147.