Primality testing

- Now we move to factoring, or irreducibility testing, of integers.

- Motivation: • natural qn. (first raised by Gauss formally).
  • commercially, appears in RSA used in browsers, file transfer applications (eg. SSH), smartcards, etc.

- The first question: Is input n prime?

Historical attempts

1) Antiquity (Eratosthenes Sieve, 300 B.C.)

$$\text{Divide } n \text{ by } 2, 3, \ldots, \lfloor \sqrt{n} \rfloor.$$

• It's doable for small n, eg. 127. But for large n, eg. $2^{127}-1$, $\sqrt{n}$ steps is way too long.

- Ideally, we want a $(\lg n)^{O(1)}$ time algorithm.

2) Fermat test (1660s).

For several a, test $a^n \equiv a \pmod{n}$.

- It is fast for a single $a \in \mathbb{Z}/n\mathbb{Z}$.
- But how many a's should we try till we can deduce "n is prime"?

- Carmichael (1910) showed the existence of composite n's s.t. $\forall a \in (\mathbb{Z}/n\mathbb{Z})^*$, $a^n \equiv a \pmod{n}$.

  e.g. $n = 561 = 3 \times 11 \times 17$.

- Alford, Granville & Pomerance (1994) showed that there are ∞ly many Carmichael numbers.

  In fact, at least $n^{2/7}$ in the set $[n]$.

3) Solovay-Strassen (1974).

- This was the first correct, "practical" primality test.

- It is based on <u>quadratic residuosity</u> and is a randomized poly-time primality test.

<u>Lemma 1</u> (Legendre symbol): For a prime p & $a \in \mathbb{Z}$, define $\left(\frac{a}{p}\right) := a^{\frac{p(p-1)}{2}} \bmod p$. Then, a is a square in $\mathbb{F}_p^*$ iff $\left(\frac{a}{p}\right) = 1$.

<u>Pf</u>:

- Seen before. $\square$

<u>Lemma 2</u> (Jacobi symbol): For numbers $a, n \in \mathbb{Z}$, define $\left(\frac{a}{n}\right) := \prod_{\text{prime } p \mid n} \left(\frac{a}{p}\right)$ (with repetition).

Then,

totally multiplicative (i) $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right) \cdot \left(\frac{b}{n}\right)$, $\forall a, b \in \mathbb{Z}$.

(ii) $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$ & $\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$, for odd $n \in \mathbb{N}$.

Gauss' (1796) quadratic reciprocity law. (iii) $\left(\frac{a}{n}\right) \cdot \left(\frac{n}{a}\right) = (-1)^{\frac{a-1}{2} \cdot \frac{n-1}{2}}$, for odd coprime $a, n \in \mathbb{N}$.

<u>Proof</u>:

- (ii) & (iii) are elementary but nontrivial.
- (iii) has more than 200 proofs known! $\square$

- Lemma 2 gives an algorithm to compute $\left(\frac{a}{n}\right)$, in a way similar to Euclid's gcd.

Algo:

1) If $(a,n) \neq 1$ then OUTPUT 0.

2.1) Replace a by $(a \bmod n) \in \left(-\frac{n}{2}, \frac{n}{2}\right]$.

2.2) If $a<0$ then reduce it to a positive case using the properties: $\left(\frac{-1}{2}\right)=1$ & $\left(\frac{a}{n}\right) = (-1)^{\frac{n-1}{2}} \cdot \left(\frac{-a}{n}\right)$.

2.3) If $2 \mid a$ then make it odd using $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$.

2.4) If $2 \mid n$ " " " " " $\left(\frac{a}{2n'}\right) = \left(\frac{a}{n'}\right)$.

2.5) If $a=1$ then OUTPUT 1.

3) OUTPUT $(-1)^{\frac{a-1}{2} \cdot \frac{n-1}{2}} \cdot \left(\frac{n}{a}\right)$.

[In each recursive step n gets at least halved. Like Euclid's gcd, the time is $\tilde{O}(\lg n)$.]

- Note that if n is prime $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \pmod n$, which may not be true in the composite n case.

- Solovay-Strassen used this idea to design a test:

Algo.: (Input: $n \in \mathbb{N}$.)

1) If $2 \mid n$ or $n = a^b$ for $b \in \mathbb{N}_{>1}$, then OUTPUT composite.
2) Pick a random $a \in [n]$.
   If $(a,n) \neq 1$ then OUTPUT composite.
3) If $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \pmod{n}$ then OUTPUT prime, else OUTPUT composite.

- We easily deduce that it runs in $\tilde{O}(\lg^2 n)$ time and that:

Claim 1: If n is prime then it outputs "prime".

Claim 2: If n is composite then $\Pr_{a \in (\mathbb{Z}/n\mathbb{Z})^*}$ [outputs "prime"] $\leq 1/2$.

Proof:

- Consider the set $B := \{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid \left(\frac{a}{n}\right) \equiv_n a^{\frac{n-1}{2}}\}$.
- It is a subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$. How big is it?

- We will later show that $B \neq (\mathbb{Z}/n\mathbb{Z})^*$.
- Thus, $|B| \leq \frac{1}{2} \cdot |(\mathbb{Z}/n\mathbb{Z})^*| = \frac{\Phi(n)}{2}$.

$$\Rightarrow \Pr_{a \in (\mathbb{Z}/n\mathbb{Z})^*} [a \in B] \leq 1/2. \qquad \square$$

<u>Connection with Riemann hypothesis</u> (RH)

- RH is a longstanding open question about the zeros of (the analytic extension of)

$$\zeta(s) := \sum_{n \geq 1} n^{-s} \quad \text{(Riemann zeta fn.)}$$

- RH has deep connections to the distribution of prime numbers.

- (Ankeny 1950 & Bach 1990) showed that:
  If $B \lneq (\mathbb{Z}/n\mathbb{Z})^*$ & GRH holds, then $\exists\, a \in \{1, \ldots, \lceil 2\lg^2 n \rceil\}$ s.t. $a \notin B$.

$\Rightarrow$ Solovay-Strassen's primality test can be derandomized, under GRH, to a deterministic poly-time test.

- Miller (1975) gave another such test, which was later made practical by Rabin (1977).

  Miller-Rabin test is the simplest & practically the most popular primality test.

- Idea: Continue beyond $a^{\frac{n-1}{2}} \pmod n$ to $a^{\frac{n-1}{4}}$, $a^{\frac{n-1}{8}}, \ldots \pmod n$. Whp we'll get a $\sqrt{1}$ other than $\pm 1 \bmod n$.

Miller-Rabin test: (Input: $n \in \mathbb{N}$ in binary.)

1) If n is even or $\exists a, b > 1$, $n = a^b$, then OUTPUT composite.

2.1) Randomly choose $a \in [n-1]$.

2.2) If $(a,n) \neq 1$ or $a^{n-1} \not\equiv 1 \pmod{n}$
then OUTPUT composite.

3) Compute k, m s.t. $n-1 = 2^k \cdot m$, odd m.

4) For $i=0$ to $(k-1)$
Compute $u_i = a^{m \cdot 2^i} \bmod n$.

5) If $\exists i$, $u_i = 1$ & $u_{i-1} \not\equiv \pm 1$ $\triangleright\ u_{i-1}^2 \equiv u_i \equiv 1$
then OUTPUT composite else OUTPUT prime.

- Its time complexity is clearly $\tilde{O}(\lg^2 n)$.

__Fact:__ If n is prime then it outputs "prime".

__Pf:__

- For prime n, $\sqrt{1}$ can only be $\pm 1 \bmod n$, since $\mathbb{Z}/n\mathbb{Z}$ is a __field__. $\square$

__Theorem:__ If n is odd & has ≥ 2 distinct prime factors then the bad a's of Miller-Rabin, i.e.

$$B := \{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid a^m = 1 \text{ or } \exists 0 \leq i \leq k, a^{m2^i} = -1\}$$

are at most $\varphi(n)/4$ many.

Proof:

- We will prove this by studying the congruences mod n via Chinese remaindering.

- Let 2^l be the highest 2-power that divides $\gcd(p-1 \mid \text{prime } p \text{ dividing } n)$.
- Define $B' := \{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid a^{m\cdot 2^{l-1}} = \pm 1\}$.

– B may not be a subgroup, but B' is.

▷ $B \subseteq B' \leq (\mathbb{Z}/n\mathbb{Z})^*$.

Pf: • Let $a \in B$.

• If $a^m = 1$ then clearly $a \in B'$.

• If $a^{m\cdot 2^i} = -1$ then $\forall p^e \| n$, $a^{m 2^i} = -1 \pmod{p^e}$.

$(\mathbb{Z}/p^e\mathbb{Z})^*$ is cyclic

$\Rightarrow 2^{i+1} \mid (p-1)$ odd

$\Rightarrow i \leq (l-1) \Rightarrow a^{m\cdot 2^{l-1}} = \pm 1 \pmod{n}$.

$\Rightarrow a \in B'$. □

▷ $\# B' = 2 \cdot \prod_{p \mid n} \left(\gcd(m, p-1) \cdot 2^{l-1}\right)$ ← distinct primes $p \mid n$

Pf:

• Let us compute $\#\{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid a^{m 2^{l-1}} = 1\}$.

• $= \prod_{p|n} \#\{a \in (\mathbb{Z}/p^{e_p}\mathbb{Z})^* \mid a^{m \cdot 2^{\ell-1}} = 1\}$

[where, $n = \prod_{p|n} p^{e_p}$ for distinct primes]

$= \prod_{p|n} \gcd(m2^{\ell-1}, \varphi(p^{e_p})) = \prod (m2^{\ell-1}, p^{e_p-1}(p-1))$

[$\because$ $(\mathbb{Z}/p^{e_p}\mathbb{Z})^*$ is a cyclic group of order $\varphi(p^{e_p})$.]

$= \prod_{p|n} (\gcd(m, p-1) \cdot 2^{\ell-1})$

[$\because$ $\varphi(p^{e_p}) = p^{e_p-1}(p-1)$; p is coprime to $2m$ & m is odd.]

• By the above count we deduce that

$\#B' = 2 \cdot \prod_{p|n} (m, p-1) \cdot 2^{\ell-1}$. Δ

$\Rightarrow \frac{\#B'}{\varphi(n)} = 2 \cdot \prod_{p|n} \frac{(m, p-1) \cdot 2^{\ell-1}}{(p-1) \cdot p^{e_p-1}}$

$< 2 \cdot \prod_{p|n} \frac{1/2}{p^{e_p-1}}$ [$\because$ the numerator divides $(p-1)/2$]

$\Rightarrow$ We are done $\begin{cases} \text{if } n \text{ has } \geq 3 \text{ prime factors, or} \\ \text{if } \exists p|n,\ e_p \geq 2. \end{cases}$

- Thus, we assume $n = p \cdot q$ for distinct primes.

$$\Rightarrow \frac{\# B'}{\varphi(n)} = 2 \cdot \frac{(p-1,m) \cdot 2^{\ell-1}}{p-1} \cdot \frac{(q-1,m) \cdot 2^{\ell-1}}{q-1}$$

$$= \frac{1}{2} \cdot \frac{(p-1,m)}{(p-1)/2^{\ell}} \cdot \frac{(q-1,m)}{(q-1)/2^{\ell}}$$

numerators divide their denominator

- RHS is $\geq 1/4$ only if
$(p-1,m) = (p-1)2^{-\ell}$ & $(q-1,m) = (q-1)2^{-\ell}$.

$\Rightarrow \exists\, p', q'$ dividing m s.t.
$p-1 = 2^{\ell} \cdot p'$ & $q-1 = 2^{\ell} \cdot q'$.

$\Rightarrow n = 2^k \cdot m + 1 = (1+2^{\ell} \cdot p') \cdot (1 + 2^{\ell} \cdot q')$

$\Rightarrow p' | q'$ & $q' | p'$

$\Rightarrow p' = q' \Rightarrow p = q$, a $\lightning$

- Thus, $\frac{\# B}{\varphi(n)} \leq \frac{\# B'}{\varphi(n)} < \frac{1}{4}$. $\square$

<u>Corollary 1:</u> Miller-Rabin test could err when n is composite, with probability $< 1/4$.

<u>Corollary 2:</u> Miller-Rabin could be derandomized,

under GRH, to a det. poly-time test.

Pf:

- We have shown that if n is composite then B' is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.

- Thus, from the "GRH connection" $\exists\ 1 \le a \le 2\lg^2 n$, $a \notin B'$.

$\Rightarrow a \notin B$, and hence Miller-Rabin works correctly with this a. $\square$

Proving Solovay-Strassen

- We now give the missing proof of the Solovay-Strassen test.

The idea is to study congruences mod n via CRT.

Theorem: If n is composite, odd & having ≥ 2 prime factors then $B := \{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid a^{\frac{n-1}{2}} \equiv (\frac{a}{n})\}$ is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.

Proof:

- Suppose $\exists$ prime p_1, $p_1^2 | n$. Let $n = p_1^{e_1} \cdots p_k^{e_k}$ for distinct primes p_i.
- Since $(\mathbb{Z}/p_1^{e_1}\mathbb{Z})^*$ is a cyclic group of order $\Phi(p_1^{e_1}) = p_1^{e_1-1} \cdot (p_1 - 1)$, we could pick its generator g.
- If $g \in B$ then $\Phi(p_1^{e_1}) \mid (n-1)$

$$\Rightarrow p_1 \mid (n-1), \text{ a } \lightning.$$

$\Rightarrow$ $g \notin B$ and we are done.

- Thus, we assume $n = p_1 \cdots p_k$. [sq-free]
- If $\exists i \in [k]$ & $g_{\in \mathbb{N}}$ s.t. $g^{\frac{n-1}{2}} \not\equiv \left(\frac{g}{p_i}\right) \pmod{p_i}$,

then we can find, by CRT, $a \equiv g \pmod{p_i}$ & $\forall j \in [k] \setminus \{i\}$, $a \equiv 1 \pmod{p_j}$.

$$\Rightarrow a^{\frac{n-1}{2}} \not\equiv \left(\frac{a}{n}\right) = \left(\frac{a}{p_i}\right) \pmod{p_i}$$

$\Rightarrow$ $a \notin B$ and we'll be done.

- Thus, the bad case is: $\forall g, \forall i, g^{\frac{n-1}{2}} \underset{p_i}{\equiv} \left(\frac{g}{p_i}\right)$.

- Now, since $k \geq 2$, we could pick an a s.t. $\left(\frac{a}{p_1}\right) = 1, \left(\frac{a}{p_2}\right) = -1$ & $a \underset{p_i}{\equiv} 1$, $\forall i \in [3 \dots k]$.

$\Rightarrow a^{\frac{n-1}{2}} \underset{p_1}{\equiv} 1, \underset{p_2}{\equiv} -1$ & $\underset{p_i}{\equiv} 1$, $\forall i \in [3 \dots k]$.

$\Rightarrow a^{\frac{n-1}{2}} \not\equiv \pm 1 \pmod{n}$

$\Rightarrow a^{\frac{n-1}{2}} \not\equiv \left(\frac{a}{n}\right) \pmod{n}$.

$\Rightarrow a \notin B$. $\square$