Primality Testing

Primes

- Natural (counting) numbers: $\mathcal{N} = \{1, 2, 3, \ldots\}$
- A number $p \in \mathcal{N}$ is called prime if it is divisible only by 1 and itself
- $p = 1$ is not considered prime; 2 is the only even prime
- Primes: 2, 3, 5, 7, 11, 13, ...
- There are infinitely many primes
- Every natural number n is factored into prime powers uniquely:

$$n = p_1^{k_1} \times p_2^{k_2} \times \cdots \times p_m^{k_m}$$

For example: $1960 = 2^3 \times 5^1 \times 7^2$

Primes

- The number of primes less than or equal to n is $\frac{n}{\log_e(n)}$

n	$n/\log_e(n)$	exact
10^2	21.7	25
10^3	144.8	168
10^6	72382.4	78498
10^9	4.8×10^7	50847534

- As we can see, primes are in abundance; we do not have scarcity
- The odds of selecting a prime is high for small numbers: if we select a 2-digit integer, the probability that it is prime is $25/100 = 25\%$
- The odds of selecting a prime less than 10^6 is $78498/10^6 \approx 7.8\%$
- If we make sure that this number is not divisible by 2 or 3, (which makes up 2/3 of integers), the odds increase to 23.5%

Primes

- As the numbers get larger, which would be the case for cryptographic applications, the ratio becomes less and less
- The ratio of 1024-bit (308-digit) primes to the 308-digit numbers is

$$\frac{1}{\log_e(2^{1024})} \approx \frac{1}{714}$$

- Therefore, if we randomly select a 308-digit integer, the probability that it is prime is $1/714$
- If we remove the multiples of 2 and 3 from this selected integer, the odds of choosing a 308-digit prime at random is improved by a factor of 3 to $1/238$

Checking for Primality vs Factoring

- Primality testing: Is $n \in \mathcal{N}$ prime?
  The answer is yes or no (we may not need the factors if n is composite)
- Factoring: What is the prime factorization of $n \in \mathcal{N}$?
  The answer is $n = p_1^{k_1} \times \cdots \times p_m^{k_m}$
- Is $2^{101} + 81 = 2535301200456458802993406410833$ prime?
  The answer: Yes
- Is $2^{101} + 71 = 2535301200456458802993406410823$ prime?
  The answer: No
- Factor $n = 2^{101} + 61 = 2535301200456458802993406410813$
  The answer: $n = 3 \times 19 \times 1201 \times 37034944570408560161757109$

Primality Testing

- The decision problem "Is n prime?" is called the primality testing
- Primality testing is easier than factorization, as might be expected, since we are not asking for the factors of n
- There are two very efficient randomized polynomial-time algorithms: **Fermat's method** and **Miller-Rabin method**
- There is also a deterministic polynomial-time algorithm invented in 2002: **The AKS algorithm**, due to three Indian computer scientists: Manindra Agrawal, Neeraj Kayal, and Nitin Saxena at the IIT Kanpur
- In the first version of their paper, time complexity was $O(b^{12})$, which was later improved to $O(b^{10.5})$ and then to $O(b^{7.5})$, where $b = \log(n)$

Fermat's Method

- Fermat's Little Theorem: If p is prime and $1 \leq a < p$, then

$$a^{p-1} = 1 \pmod{p}$$

- The contrapositive of Fermat's Little Theorem: If a and n satisfy $1 \leq a < n$ and $a^{n-1} \neq 1 \pmod{n}$, then n is composite
- Consider the list of $3^{n-1} \pmod{n}$ for $n = 4, 5, \ldots, 19$

n	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
3^{n-1}	3	1	3	1	3	0	3	1	3	1	3	9	11	1	9	1

- This shows that for all composite numbers in this range, $3^{n-1} \pmod{n}$ is distinct from 1, whereas all prime numbers satisfy $3^{n-1} = 1 \pmod{n}$

Fermat's Witness and Fermat's Liar

- Fermat's Little Theorem (and its contrapositive) provide good criteria for checking primality
- A number a in the range $a \in [1, n)$ is called a **Fermat's witness** for any $n \geq 2$, if $a^{n-1} \neq 1 \pmod{n}$
- Existence of a witness for n means n is a composite number
- A number a in the range $a \in [1, n)$ is called a **Fermat's liar** for an odd composite number $n \geq 3$, if $a^{n-1} = 1 \pmod{n}$
- Fermat's liar a is lying to us that n is prime, even though n is an odd composite number

Fermat's Witness and Fermat's Liar

- 2 is a witness for all composite n in the range $[2, 340]$ since if n is composite then $2^{n-1} \neq 1 \pmod{n}$, for $n = 2, 3, \ldots, 340$
- 2 is a liar for $n = 341$, since $2^{340} = 1 \pmod{341}$ even though it is not a prime number: $341 = 11 \cdot 31$
- 3 is a witness for 341 since $3^{340} = 56 \pmod{341}$
- Because of the existence of Fermat liars, the converse of Fermat's Little Theorem is not true: The condition that $a^{n-1} = 1 \pmod{n}$ does not imply that n is prime
- However, if n is a composite number, then there exists some Fermat's witness a for n

The Fermat Test

FERMAT(n)
Input: $n \geq 3$ is an odd integer
Step 1: Randomly choose a in the range $a \in [2, n-2]$
Step 2: $x := a^{n-1} \pmod n$
Step 2: `if` $x \neq 1 \pmod n$ `return` "n is composite"
`else return` "n is prime"

- Fermat's test is a randomized algorithm
- If the Fermat test gives the answer "n is composite", the number n is composite indeed
- However, if the Fermat test gives the answer "n is prime", the number n may or may not be prime, as there are Fermat's liars

The Fermat Test

- Consider $n = 143$ which is a composite number $143 = 11 \cdot 13$
- The table below shows Fermat's witnesses and liars for 143

Multiples of 11	11	22	33	44	55	66	77	88	99	110	121	132
Multiples of 13	13	26	39	52	65	78	91	104	117	130		
Fermat witnesses	2	3	4	5	6	7	8	9	10	14	15	16
in $\mathbb{Z}^*_{143}$	17	18	19	20	21	23	24	25	27	28	29	30
	31	32	34	35	36	37	38	40	41	42	43	45
	46	47	48	49	50	51	53	54	56	57	58	59
	60	61	62	63	64	67	68	69	70	71	72	73
	74	75	76	79	80	81	82	83	84	85	86	87
	89	90	92	93	94	95	96	97	98	100	101	102
	103	105	106	107	108	109	111	112	113	114	115	116
	118	119	120	122	123	124	125	126	127	128	129	133
	134	135	136	137	138	139	140	141				
Fermat liars	1	12	131	142								

The Fermat Test

- If we run the Fermat test on 143, the probability that it answers "n is composite" is $138/140 \approx 0.9857$, since there are only two (non trivial) Fermat liars
- In other words, the Fermat witnesses outnumber the Fermat liars clearly in this example
- If this were true for all odd composite numbers, we would have a no-biased Monte Carlo algorithm for the primality problem
- A no-biased Monte Carlo algorithm **always** gives correct "no" answers, but perhaps incorrect "yes" answers
- Unfortunately, if n is composite, the Fermat test does not say so with probability at least $1/2$ for **each given** n

Carmichael Numbers

- There exist composite numbers n for which all elements of $\mathcal{Z}_n^*$ are Fermat liars
- Such numbers are called Carmichael numbers
- The smallest Carmichael number: $561 = 3 \cdot 11 \cdot 17$
- The next 6 Carmichael numbers are 1105, 1729, 2465, 2821, 6601, 8911
- Note that Carmichael numbers have Fermat witnesses in $\mathcal{Z}_n - \mathcal{Z}_n^*$
- It was proven in 1994 by Alford, Granville, and Pomerance that there are infinitely many Carmichael numbers: Specifically they proved that there are at least $\sqrt[7]{n^2}$ Carmichael numbers between 1 and n
- Carmichael numbers have at least 3 prime factors

The Fermat Test

- Theorem: If $n \geq 3$ is an odd composite number that has at least one Fermat witness in $\mathcal{Z}_n^*$, then the Fermat test on input n gives the correct answer "n is composite" with probability at least $1/2$
- This theorem says that for many composite numbers (except Carmichael numbers) the Fermat test has a good probability bound
- The reason why the Fermat test is not a Monte Carlo algorithm for "is n prime?" problem is that $\mathcal{Z}_n^*$ contains too many Fermat liars for infinitely many numbers n, namely Carmichael numbers
- Given a Carmichael number n as input, the Fermat test gives the wrong answer "n is prime" with probability

$$\frac{\phi(n)}{n} \approx \prod(1 - \frac{1}{p}) \lesssim 1$$

The Miller-Rabin Test

MILLER-RABIN(n)

Input: $n \geq 3$ is odd, such that $n-1 = 2^k \cdot m$, for odd m

Step 1: Randomly a in the range $a \in [1, n-1]$

Step 2: $x := a^m \pmod{n}$

Step 3: `if` $x = 1 \pmod{n}$ `return` "n is prime" `and halt`

Step 4: `for` $j = 0, 1, \ldots, k-1$

Step 5: `if` $x = -1 \pmod{n}$, `return` "n is prime" `and halt`

`else` $x := x^2 \pmod{n}$

Step 6: `return` "n is composite" `and halt`

The Miller-Rabin Example

- $n = 561$ implies $n - 1 = 560 = 2^4 \cdot 35$, thus $k = 4$ and $m = 35$
- Pick $a = 2$ and compute $x := 2^{35} = 263 \pmod{561}$; $x \neq 1$
- $j = 0 \rightarrow x \neq -1 \pmod{561}$; $x := 263^2 = 166 \pmod{561}$
- $j = 1 \rightarrow x \neq -1 \pmod{561}$; $x := 166^2 = 67 \pmod{561}$
- $j = 2 \rightarrow x \neq -1 \pmod{561}$; $x := 67^2 = 1 \pmod{561}$
- $j = 3 \rightarrow x \neq -1 \pmod{561}$; $x := 1^2 = 1 \pmod{561}$
- Therefore, n is composite

Square Roots of 1 Mod n

- An element $x \in \mathcal{Z}_n$ is a quadratic residue mod n if and only if there is some $a \in [1, n)$ such that $x = a^2 \pmod{n}$
- For example, 3 is quadratic residue mod 11 since $3 = 5^2 \pmod{11}$
- If $x = 1$, then a is said to be square root of 1 mod n
- Trivially, 1 and -1 are always square roots of 1 mod m since $1^2 = 1 \pmod{n}$ and $(n-1)^2 = (-1)^2 = 1 \pmod{n}$
- The prime number 23 has 2 square roots of 1, namely 1 and 22
- The composite number $143 = 11 \cdot 13$ has 4 square roots of 1, namely 1, 12, 131, and 142

Square Roots of 1 Mod n

- Theorem: Every prime number n has only two trivial square roots of 1 mod n, namely $\pm 1 \pmod{n}$
- Hence, if n has a nontrivial (other than ± 1) square root of 1, then n must be composite
- If $n = p_1 p_2 \cdots p_k$ is composite, where $p_i > 2$ are prime numbers then the Chinese Remainder Theorem can be used to show that n has exactly 2^k square roots of 1 mod n
- The square roots of 1 mod n are all numbers $a \in [1, n)$ such that $a = \pm 1 \pmod{p_i}$ for $i = 1, 2, \ldots, k$
- Unless n has extraordinarily many prime factors, we cannot find nontrivial square roots of 1 mod n by picking random numbers a

Miller-Rabin Witnesses and Miller-Rabin Liars

- Let $n \geq 3$ be any odd number and $a \in \mathcal{Z}_n^*$
- Express $n - 1 = 2^k \cdot m$ with m is odd
- We say a is a **Miller-Rabin liar** for n if and only if n is a composite number and one of the following is **true**:
  - $a^m = 1 \pmod{n}$
  - $a^m = -1 \pmod{n}$
  - $a^{2m} = -1 \pmod{n}$
  - $a^{2^2 m} = -1 \pmod{n}$
  - $\cdots$
  - $a^{2^{k-1} m} = -1 \pmod{n}$
- We say a is a **Miller-Rabin witness** for n if and only if a is not a Miller-Rabin liar

Miller-Rabin Witnesses and Miller-Rabin Liars

- Consider the Carmichael number $n = 561 = 3 \cdot 11 \cdot 17$
- We have $n - 1 = 560 = 2^4 \cdot 35$, and thus $k = 4$ and $m = 35$
- By enumeration, we show that 561 has 10 Miller-Rabin liars

a	a^{35}	a^{70}	a^{140}	a^{280}	a^{560}
1	1	1	1	1	1
50	-1	1	1	1	1
101	-1	1	1	1	1
103	1	1	1	1	1
256	1	1	1	1	1
305	-1	1	1	1	1
458	-1	1	1	1	1
460	1	1	1	1	1
511	1	1	1	1	1
560	-1	1	1	1	1

The rest of numbers in $\mathcal{Z}_{561}^*$ are all Miller-Rabin witnesses

The Miller-Rabin Test

- Theorem: If there exists a Miller-Rabin witness for n, then n is composite
- Theorem: If $n \geq 3$ is an odd composite number, then there are at most $\frac{n-1}{4}$ Miller-Rabin liars
- Theorem: The Miller-Rabin Test has an error probability of at most $1/4$
- The Miller-Rabin test is very efficient and has a very good probability bound — it is the preferred algorithm for generating large primes used in the RSA algorithm, the Diffie-Hellman key exchange algorithm, or any of the public-key cryptographic protocols where large primes are needed
- There is another probabilistic algorithm for primality testing, called Solovay-Strassen test, however, it is less efficient and less accurate, and therefore, less popular