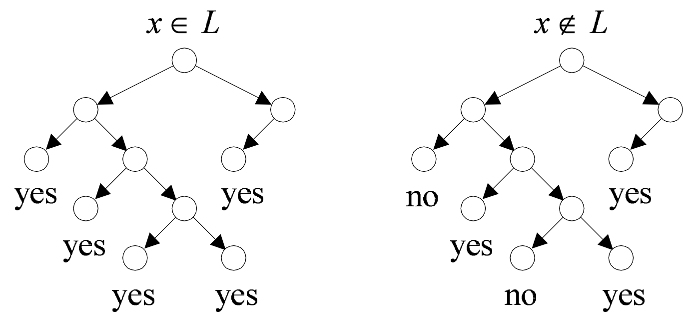
coNP and Function Problems

coNP

- By definition, coNP is the class of problems whose complement is in NP.
- NP is the class of problems that have succinct certificates (recall Proposition 31 on p. 254).
- coNP is therefore the class of problems that have succinct disqualifications:
  - A "no" instance of a problem in coNP possesses a short proof of its being a "no" instance.
  - Only "no" instances have such proofs.

coNP (continued)

- Suppose L is a coNP problem.
- There exists a polynomial-time nondeterministic algorithm M such that:
  - If $x \in L$, then $M(x) =$ "yes" for all computation paths.
  - If $x \notin L$, then $M(x) =$ "no" for some computation path.



coNP (concluded)

- Clearly $\mathrm{P} \subseteq \mathrm{coNP}$.
- It is not known if

$$\mathrm{P} = \mathrm{NP} \cap \mathrm{coNP}.$$

  - Contrast this with

$$\mathrm{R} = \mathrm{RE} \cap \mathrm{coRE}$$

    (see Proposition 11 on p. 126).

Some coNP Problems

- VALIDITY $\in$ coNP.
  - If ϕ is not valid, it can be disqualified very succinctly: a truth assignment that does not satisfy it.
- SAT COMPLEMENT $\in$ coNP.
  - The disqualification is a truth assignment that satisfies it.
- HAMILTONIAN PATH COMPLEMENT $\in$ coNP.
  - The disqualification is a Hamiltonian path.

An Alternative Characterization of coNP

Proposition 43 *Let $L \subseteq \Sigma^*$ be a language. Then $L \in$ coNP if and only if there is a polynomially decidable and polynomially balanced relation R such that*

$$L = \{x : \forall y\, (x, y) \in R\}.$$

(As on p. 253, we assume $|\,y\,| \leq |\,x\,|^k$ for some k.)

- $\bar{L} = \{x : (x, y) \in \neg R \text{ for some } y\}$.
- Because $\neg R$ remains polynomially balanced, $\bar{L} \in \mathrm{NP}$ by Proposition 31 (p. 254).
- Hence $L \in$ coNP by definition.

coNP Completeness

Proposition 44 *L is NP-complete if and only if its complement $\bar{L} = \Sigma^* - L$ is coNP-complete.*

Proof ($\Rightarrow$; the $\Leftarrow$ part is symmetric)

- Let $\bar{L}'$ be any coNP language.
- Hence $L' \in \mathrm{NP}$.
- Let R be the reduction from L' to L.
- So $x \in L'$ if and only if $R(x) \in L$.
- So $x \in \bar{L}'$ if and only if $R(x) \in \bar{L}$.
- R is a reduction from $\bar{L}'$ to $\bar{L}$.

Some coNP-Complete Problems

- SAT COMPLEMENT is coNP-complete.
  - SAT COMPLEMENT is the complement of SAT.
- VALIDITY is coNP-complete.
  - ϕ is valid if and only if $\neg\phi$ is not satisfiable.
  - The reduction from SAT COMPLEMENT to VALIDITY is hence easy.
- HAMILTONIAN PATH COMPLEMENT is coNP-complete.

Possible Relations between P, NP, coNP

1. P = NP = coNP.
2. NP = coNP but P $\neq$ NP.
3. NP $\neq$ coNP and P $\neq$ NP.
   - This is current "consensus."

coNP Hardness and NP Hardness[a]

Proposition 45 *If a coNP-hard problem is in NP, then NP = coNP.*

- Let $L \in$ NP be coNP-hard.
- Let NTM M decide L.
- For any $L' \in$ coNP, there is a reduction R from L' to L.
- $L' \in$ NP as it is decided by NTM $M(R(x))$.
  - Alternatively, NP is closed under complement.
- Hence coNP $\subseteq$ NP.
- The other direction NP $\subseteq$ coNP is symmetric.

[a]Brassard (1979); Selman (1978).

coNP Hardness and NP Hardness (concluded)

Similarly,

Proposition 46 *If an NP-hard problem is in coNP, then NP = coNP.*

Hence NP-complete problems are unlikely to be in coNP and coNP-complete problems are unlikely to be in NP.

The Primality Problem

- An integer p is **prime** if $p > 1$ and all positive numbers other than 1 and p itself cannot divide it.
- PRIMES asks if an integer N is a prime number.
- Dividing N by $2, 3, \ldots, \sqrt{N}$ is *not* efficient.
  - The length of N is only $\log N$, but $\sqrt{N} = 2^{0.5 \log N}$.
- A polynomial-time algorithm for PRIMES was not found until 2002 by Agrawal, Kayal, and Saxena!
- We will focus on efficient "probabilistic" algorithms for PRIMES (used in *Mathematica*, e.g.).


```
if n = a^b for some a, b > 1 then
  return "composite";
end if
for r = 2, 3, ..., n − 1 do
  if gcd(n, r) > 1 then
    return "composite";
  end if
  if r is a prime then
    Let q be the largest prime factor of r − 1;
    if q ≥ 4√r log n and n^((r−1)/q) ≠ 1 mod r then
      break; {Exit the for-loop.}
    end if
  end if
end for{r − 1 has a prime factor q ≥ 4√r log n.}
for a = 1, 2, ..., 2√r log n do
  if (x − a)^n ≠ (x^n − a) mod (x^r − 1) in Z_n[x] then
    return "composite";
  end if
end for
return "prime"; {The only place with "prime" output.}
```


DP

- DP $\equiv$ NP $\cap$ coNP is the class of problems that have succinct certificates and succinct disqualifications.
  - Each "yes" instance has a succinct certificate.
  - Each "no" instance has a succinct disqualification.
  - No instances have both.
- P $\subseteq$ DP.
- We will see that PRIMES $\in$ DP.
  - In fact, PRIMES $\in$ P as mentioned earlier.

Primitive Roots in Finite Fields

Theorem 47 (Lucas and Lehmer (1927)) [a] *A number $p > 1$ is prime if and only if there is a number $1 < r < p$ (called the* **primitive root** *or* **generator***) such that*

1. $r^{p-1} = 1 \bmod p$, *and*
2. $r^{(p-1)/q} \neq 1 \bmod p$ *for all prime divisors q of $p - 1$.*

- We will prove the theorem later.

---

[a] François Edouard Anatole Lucas (1842–1891); Derrick Henry Lehmer (1905–1991).

Pratt's Theorem

Theorem 48 (Pratt (1975)) *PRIMES $\in NP \cap coNP$.*

- PRIMES is in coNP because a succinct disqualification is a divisor.
- Suppose p is a prime.
- p's certificate includes the r in Theorem 47 (p. 346).
- Use recursive doubling to check if $r^{p-1} = 1 \bmod p$ in time polynomial in the length of the input, $\log_2 p$.
- We also need all *prime* divisors of $p-1$: $q_1, q_2, \ldots, q_k$.
- Checking $r^{(p-1)/q_i} \neq 1 \bmod p$ is also easy.

The Proof (concluded)

- Checking $q_1, q_2, \ldots, q_k$ are all the divisors of $p-1$ is easy.
- We still need certificates for the primality of the q_i's.
- The complete certificate is recursive and tree-like:

$$C(p) = (r; q_1, C(q_1), q_2, C(q_2), \ldots, q_k, C(q_k)).$$

- $C(p)$ can also be checked in polynomial time.
- We next prove that $C(p)$ is succinct.

The Succinctness of the Certificate

Lemma 49 *The length of $C(p)$ is at most quadratic at $5\log_2^2 p$.*

- This claim holds when $p = 2$ or $p = 3$.
- In general, $p-1$ has $k < \log_2 p$ prime divisors $q_1 = 2, q_2, \ldots, q_k$.
- $C(p)$ requires: 2 parentheses and $2k < 2\log_2 p$ separators (length at most $2\log_2 p$ long), r (length at most $\log_2 p$), $q_1 = 2$ and its certificate 1 (length at most 5 bits), the q_i's (length at most $2\log_2 p$), and the $C(q_i)$s.

The Proof (concluded)

- $C(p)$ is succinct because

$$\begin{aligned}
|C(p)| &\leq 5\log_2 p + 5 + 5\sum_{i=2}^{k} \log_2^2 q_i \\
&\leq 5\log_2 p + 5 + 5\left(\sum_{i=2}^{k} \log_2 q_i\right)^2 \\
&\leq 5\log_2 p + 5 + 5\log_2^2 \frac{p-1}{2} \\
&< 5\log_2 p + 5 + 5(\log_2 p - 1)^2 \\
&= 5\log_2^2 p + 10 - 5\log_2 p \leq 5\log_2^2 p
\end{aligned}$$

for $p \geq 4$.

Basic Modular Arithmetics[a]

- Let $m, n \in \mathbb{Z}^+$.
- $m|n$ means m divides n and m is n's **divisor**.
- We call the numbers $0, 1, \ldots, n-1$ the **residue** modulo n.
- The **greatest common divisor** of m and n is denoted $\gcd(m, n)$.
- The r in Theorem 47 (p. 346) is a primitive root of p.
- We now prove the existence of primitive roots and then Theorem 47.

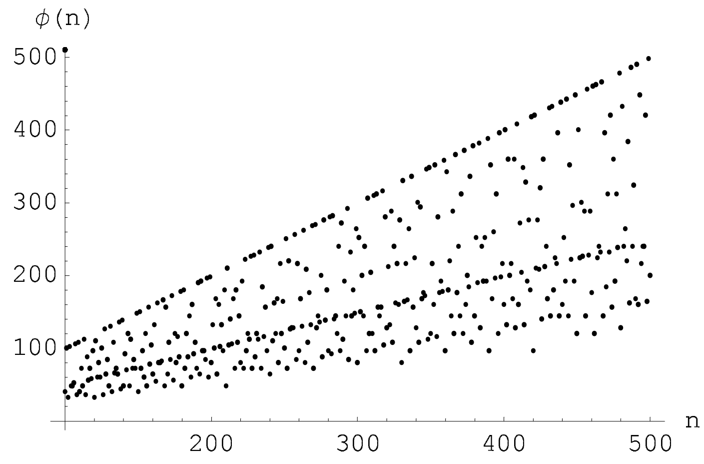
[a]Carl Friedrich Gauss.

Euler's[a] Totient or Phi Function

- Let
$$\Phi(n) = \{m : 1 \leq m < n, \gcd(m, n) = 1\}$$
be the set of all positive integers less than n that are prime to n (Z_n^* is a more popular notation).
  - $\Phi(12) = \{1, 5, 7, 11\}$.
- Define **Euler's function** of n to be $\phi(n) = |\Phi(n)|$.
- $\phi(p) = p - 1$ for prime p, and $\phi(1) = 1$ by convention.
- Euler's function is not expected to be easy to compute without knowing n's factorization.

[a]Leonhard Euler (1707–1783).

eulerphi.nb



Two Properties of Euler's Function

The inclusion-exclusion principle[a] can be used to prove the following.

Lemma 50 $\phi(n) = n \prod_{p|n} (1 - \frac{1}{p})$.

- If $n = p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$ is the prime factorization of n, then
$$\phi(n) = n \prod_{i=1}^{t} \left(1 - \frac{1}{p_i}\right).$$

Corollary 51 $\phi(mn) = \phi(m)\,\phi(n)$ *if* $\gcd(m, n) = 1$.

[a]See my *Discrete Mathematics* lecture notes.

A Key Lemma

Lemma 52 $\sum_{m|n} \phi(m) = n$.

- Let $\prod_{i=1}^{\ell} p_i^{k_i}$ be the prime factorization of n and consider

$$\prod_{i=1}^{\ell} [\, \phi(1) + \phi(p_i) + \cdots + \phi(p_i^{k_i})\,]. \tag{4}$$

- Equation (4) equals n because $\phi(p_i^k) = p_i^k - p_i^{k-1}$ by Lemma 50.
- Expand Eq. (4) to yield $\sum_{k_1' \le k_1, \ldots, k_\ell' \le k_\ell} \prod_{i=1}^{\ell} \phi(p_i^{k_i'})$.

The Proof (concluded)

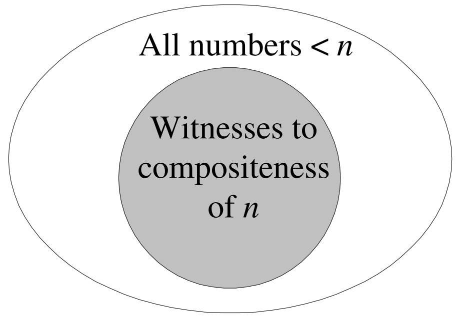
- By Corollary 51 (p. 354),

$$\prod_{i=1}^{\ell} \phi(p_i^{k_i'}) = \phi\left(\prod_{i=1}^{\ell} p_i^{k_i'} \right).$$

- Each $\prod_{i=1}^{\ell} p_i^{k_i'}$ is a unique divisor of $n = \prod_{i=1}^{\ell} p_i^{k_i}$.
- Equation (4) becomes

$$\sum_{m|n} \phi(m).$$

The Density Attack for PRIMES



- It works, but does it work well?

Factorization and Euler's Function

- The ratio of numbers $\le n$ relatively prime to n is $\phi(n)/n$.
- When $n = pq$, where p and q are distinct primes,

$$\frac{\phi(n)}{n} = \frac{pq - p - q + 1}{pq} > 1 - \frac{1}{q} - \frac{1}{p}.$$

  - The "density attack" to factor $n = pq$ hence takes $\Omega(\sqrt{n})$ steps on average when $p \sim q = O(\sqrt{n}\,)$.
  - This running time is exponential: $\Omega(2^{0.5 \log_2 n})$.

The Chinese Remainder Theorem

- Let $n = n_1 n_2 \cdots n_k$, where n_i are pairwise relatively prime.
- For any integers $a_1, a_2, \ldots, a_k$, the set of simultaneous equations

$$\begin{aligned} x &= a_1 \bmod n_1, \\ x &= a_2 \bmod n_2, \\ &\vdots \\ x &= a_k \bmod n_k, \end{aligned}$$

has a unique solution modulo n for the unknown x.

Fermat's "Little" Theorem[a]

Lemma 53 *For all* $0 < a < p$, $a^{p-1} = 1 \bmod p$.

- Consider $a\Phi(p) = \{am \bmod p : m \in \Phi(p)\}$.
- $a\Phi(p) = \Phi(p)$.
  - Suppose $am = am' \bmod p$ for $m > m'$, where $m, m' \in \Phi(p)$.
  - That means $a(m - m') = 0 \bmod p$, and p divides a or $m - m'$, which is impossible.
- Hence $(p-1)! = a^{p-1}(p-1)! \bmod p$.
- Finally, $a^{p-1} = 1 \bmod p$ because $p \not| (p-1)!$.

[a] Pierre de Fermat (1601–1665).

The Fermat-Euler Theorem

Corollary 54 *For all* $a \in \Phi(n)$, $a^{\phi(n)} = 1 \bmod n$.

- As $12 = 2^2 \times 3$,
$$\phi(12) = 12 \times \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) = 4$$
- In fact, $\Phi(12) = \{1, 5, 7, 11\}$.
- For example,
$$5^4 = 625 = 1 \bmod 12.$$

Exponents

- The **exponent** of $m \in \Phi(p)$ is the least $k \in \mathbb{Z}^+$ such that
$$m^k = 1 \bmod p.$$
- Every residue $s \in \Phi(p)$ has an exponent.
  - $1, s, s^2, s^3, \ldots$ eventually repeats itself, say $s^i = s^j \bmod p$, which means $s^{j-i} = 1 \bmod p$.
- If the exponent of m is k and $m^\ell = 1 \bmod p$, then $k | \ell$.
  - Otherwise, $\ell = qk + a$ for $0 < a < k$, and $m^\ell = m^{qk+a} = m^a = 1 \bmod p$, a contradiction.

Lemma 55 *Any nonzero polynomial of degree* k *has at most* k *distinct roots modulo* p.

Exponents and Primitive Roots

- From Fermat's "little" theorem, all exponents divide $p-1$.
- A primitive root of p is thus a number with exponent $p-1$.
- Let $R(k)$ denote the total number of residues in $\Phi(p)$ that have exponent k.
- We already knew that $R(k) = 0$ for $k \nmid (p-1)$.
- So $\sum_{k|(p-1)} R(k) = p-1$ as every number has an exponent.

Size of $R(k)$

- Any $a \in \Phi(p)$ of exponent k satisfies $x^k = 1 \bmod p$.
- Hence there are at most k residues of exponent k, i.e., $R(k) \leq k$, by Lemma 55 on p. 362.
- Let s be a residue of exponent k.
- $1, s, s^2, \ldots, s^{k-1}$ are all distinct modulo p.
  - Otherwise, $s^i = s^j \bmod p$ with $i < j$ and s is of exponent $j - i < k$, a contradiction.
- As all these k distinct numbers satisfy $x^k = 1 \bmod p$, they are all the solutions of $x^k = 1 \bmod p$.
- But do all of them have exponent k (i.e., $R(k) = k$)?

Size of $R(k)$ (continued)

- And if not (i.e., $R(k) < k$), how many of them do?
- Suppose $\ell < k$ and $\ell \notin \Phi(k)$ with $\gcd(\ell, k) = d > 1$.
- Then
$$(s^\ell)^{k/d} = 1 \bmod p.$$
- Therefore, s^ℓ has exponent at most k/d, which is less than k.
- We conclude that
$$R(k) \leq \phi(k).$$

Size of $R(k)$ (concluded)

- Because all $p-1$ residues have an exponent,
$$p - 1 = \sum_{k|(p-1)} R(k) \leq \sum_{k|(p-1)} \phi(k) = p - 1$$
by Lemma 51 on p. 354.
- Hence
$$R(k) = \begin{cases} \phi(k) & \text{when } k|(p-1) \\ 0 & \text{otherwise} \end{cases}$$
- In particular, $R(p-1) = \phi(p-1) > 0$, and p has at least one primitive root.
- This proves one direction of Theorem 47 (p. 346).