Randomized Algorithms[a]

- Randomized algorithms flip unbiased coins.
- There are important problems for which there are no known efficient *deterministic* algorithms but for which very efficient randomized algorithms exist.
  - Extraction of square roots, for instance.
- There are problems where randomization is *necessary*.
  - Secure protocols.
- Randomized version can be more efficient.
  - Parallel algorithm for maximal independent set.
- Are randomized algorithms algorithms?

[a] Rabin (1976), Solovay and Strassen (1977).

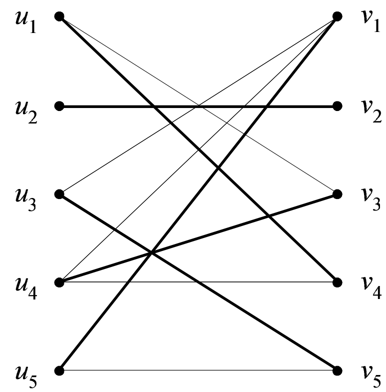
Bipartite Perfect Matching

- We are given a **bipartite graph** $G = (U, V, E)$.
  - $U = \{u_1, u_2, \ldots, u_n\}$.
  - $V = \{v_1, v_2, \ldots, v_n\}$.
  - $E \subseteq U \times V$.
- We are asked if there is a **perfect matching**.
  - A permutation π of $\{1, 2, \ldots, n\}$ such that

$$(u_i, v_{\pi(i)}) \in E$$

for all $u_i \in U$.

A Perfect Matching



Symbolic Determinants

- Given a bipartite graph G, construct the $n \times n$ matrix A^G whose (i, j)th entry A^G_{ij} is a variable x_{ij} if $(u_i, v_j) \in E$ and zero otherwise.
- The **determinant** of A^G is

$$\det(A^G) = \sum_{\pi} \sigma(\pi) \prod_{i=1}^{n} A^G_{i,\pi(i)}. \tag{5}$$

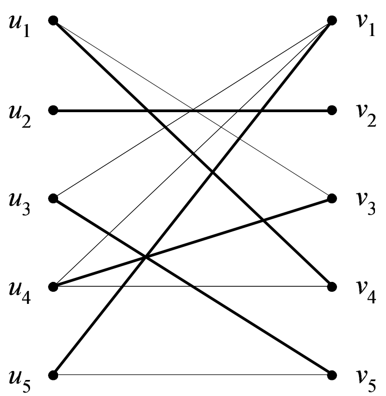
  - π ranges over all permutations of n elements.
  - $\sigma(\pi)$ is 1 if π is the product of an even number of transpositions and -1 otherwise.

Determinant and Bipartite Perfect Matching

- In $\sum_\pi \sigma(\pi) \prod_{i=1}^n A^G_{i,\pi(i)}$, note the following:
  - Each summand corresponds to a possible prefect matching π.
  - As all variables appear only *once*, all of these summands are different monomials and will not cancel.
- It is essentially an exhaustive enumeration.

Proposition 56 (Edmonds (1967)) *G has a perfect matching if and only if* $\det(A^G)$ *is not identically zero.*

A Perfect Matching in a Bipartite Graph



The Perfect Matching in the Determinant

- The matrix is

$$A^G = \begin{bmatrix} 0 & 0 & x_{13} & \boxed{x_{14}} & 0 \\ 0 & \boxed{x_{22}} & 0 & 0 & 0 \\ x_{31} & 0 & 0 & 0 & \boxed{x_{35}} \\ x_{41} & 0 & \boxed{x_{43}} & x_{44} & 0 \\ \boxed{x_{51}} & 0 & 0 & 0 & x_{55} \end{bmatrix}.$$

- $\det(A^G) = -x_{14}x_{22}x_{35}x_{43}x_{51} + x_{13}x_{22}x_{35}x_{44}x_{51} + x_{14}x_{22}x_{31}x_{43}x_{55} - x_{13}x_{22}x_{31}x_{44}x_{55}$, each denoting a perfect matching.

How To Test If a Polynomial Is Identically Zero?

- $\det(A^G)$ is a polynomial in n^2 variables.
- There are exponentially many terms in $\det(A^G)$.
- Expanding the determinant polynomial is not feasible.
  - Too many terms.
- Observation: If $\det(A^G)$ is *identically zero*, then it remains zero if we substitute *arbitrary* integers for the variables $x_{11}, \ldots, x_{nn}$.
- What is the likelihood of obtaining a zero when $\det(A^G)$ is *not* identically zero?

Number of Roots of a Polynomials

Lemma 57 (Schwartz (1980)) *Let* $p(x_1, x_2, \ldots, x_m) \not\equiv 0$ *be a polynomial in* m *variables each of degree at most* d*. Let* $M \in \mathbb{Z}^+$*. Then the number of* m*-tuples*

$$(x_1, x_2, \ldots, x_m) \in \{0, 1, \ldots, M-1\}^m$$

such that $p(x_1, x_2, \ldots, x_m) = 0$ *is*

$$\leq mdM^{m-1}.$$

- By induction on m.

Density Attack

- The density of roots in the domain is at most

$$\frac{mdM^{m-1}}{M^m} = \frac{md}{M}.$$

- A sampling algorithm to test if $p(x_1, x_2, \ldots, x_m) \not\equiv 0$.

Choose $i_1, \ldots, i_m$ from $\{0, 1, \ldots, M-1\}$ randomly;
if $p(i_1, i_2, \ldots, i_m) \neq 0$ **then**
  return "p is not identically zero";
else
  return "p is identically zero";
end if

A Randomized Bipartite Perfect Matching Algorithm[a]

Choose n^2 integers $i_{11}, \ldots, i_{nn}$ from $\{0, 1, \ldots, b-1\}$ randomly;
Calculate $\det(A^G(i_{11}, \ldots, i_{nn}))$ by Gaussian elimination;
if $\det(A^G(i_{11}, \ldots, i_{nn})) \neq 0$ **then**
  return "G has a perfect matching";
else
  return "G has no perfect matchings";
end if

[a]Lovász (1979).

Analysis

- Pick $b = 2n^2$.
- If G has no perfect matchings, the algorithm will always be correct.
- Suppose G has a perfect matching.
  - The algorithm will answer incorrectly with probability at most $n^2 d/b = 0.5$ because $d = 1$.
  - Run the algorithm *independently* k times and output "G has no perfect matchings" if they all say no.
  - The error probability is now reduced to at most 2^{-k}.

Monte Carlo Algorithms[a]

- The randomized bipartite perfect matching algorithm is called a **Monte Carlo algorithm** in the sense that
  - If the algorithm finds that a matching exists, it is always correct (no **false positives**).
  - If the algorithm answers in the negative, then it may make an error (**false negative**).
- The algorithm makes a false negative with probability ≤ 0.5.
- This probability is *not* over the space of all graphs or determinants, but *over* the algorithm's own coin flips.
  - It holds for *any* bipartite graph.

[a]Metropolis and Ulam (1949).

The Markov Inequality[a]

Lemma 58 *Let x be a random variable taking nonnegative integer values. Then for any $k > 0$,*

$$\text{prob}[x \geq kE[x]] \leq 1/k.$$

- Let p_i denote the probability that $x = i$.

$$\begin{aligned} E[x] &= \sum_i ip_i \\ &= \sum_{i<kE[x]} ip_i + \sum_{i \geq kE[x]} ip_i \\ &\geq kE[x] \times \text{prob}[x \geq kE[x]]. \end{aligned}$$

[a]Andrei Andreyevich Markov (1856–1922).

An Application of Markov's Inequality

- Algorithm C runs in expected time $T(n)$ and always gives the right answer.
- Consider an algorithm that runs C for time $kT(n)$ and rejects the input if C does not stop within the time bound.
- By Markov's inequality, this new algorithm runs in time $kT(n)$ and gives the wrong answer with probability $\leq 1/k$.
- By running this algorithm m times, we reduce the error probability to $\leq k^{-m}$.

An Application of Markov's Inequality (concluded)

- Suppose, instead, we run the algorithm for the same running time $mkT(n)$ and rejects the input if it does not stop within the time bound.
- By Markov's inequality, this new algorithm gives the wrong answer with probability $\leq 1/(mk)$.
- This is a far cry from the previous algorithm's error probability of $\leq k^{-m}$.
- The loss comes from the fact that Markov's inequality does not take advantage of any specific feature of the random variable.

Primality Tests

- PRIMES asks if a number N is a prime.
- The classic algorithm tests if $k \mid N$ for $k = 2, 3, \ldots, \sqrt{N}$.
- But it runs in $\Omega(2^{n/2})$ steps, where $n = |N| = \log_2 N$.

The Density Attack for PRIMES

```
Pick k ∈ {2, ..., N − 1} randomly; {Assume N > 2.}
if k | N then
    return "N is a composite";
else
    return "N is a prime";
end if
```


Analysis

- Suppose $N = PQ$, a product of 2 primes.
- The probability of success is

$$< 1 - \frac{\phi(N)}{N} = 1 - \frac{(P-1)(Q-1)}{PQ} = \frac{P+Q-1}{PQ}.$$

- In the case where $P \approx Q$, this probability becomes

$$< \frac{1}{P} + \frac{1}{Q} \approx \frac{2}{\sqrt{N}}.$$

- This probability is exponentially small.

The Fermat Test for Primality

- Fermat's "little" theorem on p. 341 suggests the following primality test for any given number p:
  - Pick a number a randomly from $\{1, 2, \ldots, N-1\}$.
  - If $a^{N-1} \neq 1 \bmod N$, then declare "N is composite."
  - Otherwise, declare "N is probably prime."
- Unfortunately, there are composite numbers called **Carmichael numbers** that will pass the Fermat test for *all* $a \in \{1, 2, \ldots, N-1\}$.
- There are infinitely many Carmichael numbers.[a]

[a] Alford, Granville, and Pomerance (1992).

Square Roots Modulo a Prime

- Equation $x^2 = a \bmod p$ has at most two (distinct) roots by Lemma 55 on p. 343.
  - The roots are called **square roots**.
  - Numbers a with square roots and $\gcd(a, p) = 1$ are called **quadratic residues**.
    - They are $1^2 \bmod p, 2^2 \bmod p, \ldots, (p-1)^2 \bmod p$.
- We shall show that a number either has two roots or has none, and testing which is true is trivial.
- But there are no known efficient *deterministic* algorithms to find the roots.

Euler's Test

Lemma 59 (Euler) *Let p be an odd prime and $a \neq 0 \bmod p$.*

1. *If $a^{(p-1)/2} = 1 \bmod p$, then $x^2 = a \bmod p$ has two roots.*
2. *If $a^{(p-1)/2} \neq 1 \bmod p$, then $a^{(p-1)/2} = -1 \bmod p$ and $x^2 = a \bmod p$ has no roots.*

- Let r be a primitive root of p.
- By Fermat's "little" theorem, $r^{(p-1)/2}$ is a square root of 1, so $r^{(p-1)/2} = \pm 1 \bmod p$.
- But as r is a primitive root, $r^{(p-1)/2} = -1 \bmod p$.

The Proof (concluded)

- If $a = r^{2j}$, then $a^{(p-1)/2} = r^{j(p-1)} = 1 \bmod p$ and its two distinct roots are $r^j, -r^j (= r^{j+(p-1)/2})$.
- Since there are $(p-1)/2$ such a's, and each such a has two distinct roots, we have run out of *square roots*.
  - $\{c : c^2 = a \bmod p\} = \{1, 2, \ldots, p-1\}$.
- If $a = r^{2j+1}$, then it has no roots because all the square roots have taken.
- $a^{(p-1)/2} = [\, r^{(p-1)/2} \,]^{2j+1} = (-1)^{2j+1} = -1 \bmod p$.

The Legendre Symbol[a] and Quadratic Residuacity Test

- So $a^{(p-1)/2} \bmod p = \pm 1$ for $a \neq 0 \bmod p$.
- For odd prime p, define the **Legendre symbol** $(a \,|\, p)$ as

$$(a \,|\, p) = \begin{cases} 0 & \text{if } p \,|\, a \\ 1 & \text{if } a \text{ is a quadratic residue modulo } p \\ -1 & \text{if } a \text{ is a \textbf{quadratic nonresidue} modulo } p \end{cases}$$

- Euler's test implies $a^{(p-1)/2} = (a \,|\, p) \bmod p$ for any odd prime p and any integer a.
- Note that $(ab|p) = (a|p)(b|p)$.

[a] Andrien-Marie Legendre (1752–1833).

Gauss's Lemma

Lemma 60 (Gauss) *Let p and q be two odd primes. Then $(q|p) = (-1)^m$, where m is the number of residues in $R = \{iq \bmod p : 1 \leq i \leq (p-1)/2\}$ that are greater than $(p-1)/2$.*

- All residues in R are distinct.
  - If $iq = jq \bmod p$, then $p|(j-i)\, q$ or $p|q$.
- No two elements of R add up to p.
  - If $iq + jq = 0 \bmod p$, then $p|(i+j)\, q$ or $p|q$.
- Consider the set R' of residues that result from R if we replace each of the m elements $a \in R$, where $a > (p-1)/2$, by $p - a$.

The Proof (concluded)

- All residues in R' are now at most $(p-1)/2$.
- In fact, $R' = \{1, 2, \ldots, (p-1)/2\}$.
  - Otherwise, two elements of R would add up to p.
- Alternatively, $R' = \{\pm iq \bmod p : 1 \leq i \leq (p-1)/2\}$, where exactly m of the elements have the minus sign.
- Take the product of all elements in the two representations of R'.
- So $[(p-1)/2]! = (-1)^m q^{(p-1)/2} [(p-1)/2]! \bmod p$.
- Because $\gcd([(p-1)/2]!, p) = 1$, the lemma follows.

Legendre's Law of Quadratic Reciprocity[a]

- Let p and q be two odd primes.
- Then their Legendre symbols are identical unless both numbers are 3 mod 4.

Lemma 61 (Legendre (1785), Gauss) $(p|q)(q|p) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$.

- Sum the elements of R' in the previous proof in mod2.
- On one hand, this is just $\sum_{i=1}^{(p-1)/2} i \bmod 2$.

[a] First stated by Euler in 1751. Legendre (1785) did not give a correct proof. Gauss proved the theorem when he was 19. He gave at least 6 different proofs during his life. The 152nd proof appeared in 1963.

The Proof (continued)

- On the other hand, the sum equals

$$\sum_{i=1}^{(p-1)/2} \left(qi - p \left\lfloor \frac{iq}{p} \right\rfloor \right) + mp \bmod 2$$

$$= \left(q \sum_{i=1}^{(p-1)/2} i - p \sum_{i=1}^{(p-1)/2} \left\lfloor \frac{iq}{p} \right\rfloor \right) + mp \bmod 2.$$

  - Signs are irrelevant under mod2.
  - m is as in Lemma 60 (p. 383).

The Proof (continued)

- After ignoring odd multipliers and noting that the first term above equals $\sum_{i=1}^{(p-1)/2} i$:

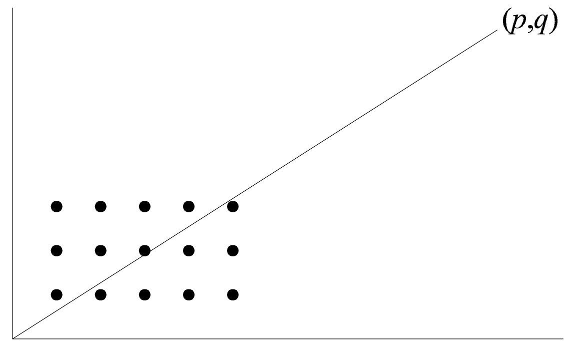
$$m = \sum_{i=1}^{(p-1)/2} \left\lfloor \frac{iq}{p} \right\rfloor \bmod 2.$$

- $\sum_{i=1}^{(p-1)/2} \lfloor \frac{iq}{p} \rfloor$ is the number of positive integral points in the $\frac{p-1}{2} \times \frac{q-1}{2}$ rectangle that are under the line between $(0,0)$ and the point (p,q).

The Proof (concluded)

- From Gauss's lemma on p. 383, $(q|p)$ is $(-1)^m$.
- Repeat the proof with p and q reversed.
- We obtain $(p|q)$ is -1 raised to the number of positive integral points in the $\frac{p-1}{2} \times \frac{q-1}{2}$ rectangle that are above the line between $(0,0)$ and the point (p,q).
- So $(p|q)(q|p)$ is -1 raised to the total number of integral points in the $\frac{p-1}{2} \times \frac{q-1}{2}$ rectangle, which is $\frac{p-1}{2} \frac{q-1}{2}$.

Eisenstein's Rectangle



$p = 11$ and $q = 7$.

The Jacobi Symbol[a]

- The Legendre symbol only works for odd *prime* moduli.
- The **Jacobi symbol** $(a \,|\, m)$ extends it to cases where m is not prime.
- Let $m = p_1 p_2 \cdots p_k$ be the prime factorization of m.
- When $m > 1$ is odd and $\gcd(a,m) = 1$, then

$$(a|m) = \prod_{i=1}^{k} (a \,|\, p_i).$$

- Define $(a \,|\, 1) = 1$.

[a]Carl Jacobi (1804–1851).

Properties of the Jacobi Symbol

The Jacobi symbol has the following properties, for arguments for which it is defined.

1. $(ab \mid m) = (a \mid m)(b \mid m)$.
2. $(a \mid m_1 m_2) = (a \mid m_1)(a \mid m_2)$.
3. If $a = b \bmod m$, then $(a \mid m) = (b \mid m)$.
4. $(-1 \mid m) = (-1)^{(m-1)/2}$ (by Lemma 60 on p. 383).
5. $(2 \mid m) = (-1)^{(m^2-1)/8}$ (by Lemma 60 on p. 383).
6. If a and m are both odd, then $(a \mid m)(m \mid a) = (-1)^{(a-1)(m-1)/4}$.

Calculation of (2200|999)

Similar to the Euclidean algorithm and does *not* require factorization.

$$\begin{aligned}
& (202|999) = (-1)^{(999^2-1)/8}(101|999) \\
= \; & (-1)^{124750}(101|999) = (101|999) \\
= \; & (-1)^{(100)(998)/4}(999|101) = (-1)^{24950}(999|101) \\
= \; & (999|101) = (90|101) = (-1)^{(101^2-1)/8}(45|101) \\
= \; & (-1)^{1275}(45|101) = -(45|101) \\
= \; & -(-1)^{(44)(100)/4}(101|45) = -(101|45) = -(11|45) \\
= \; & -(-1)^{(10)(44)/4}(45|11) = -(45|11) \\
= \; & -(1|11) = -(11|1) = -1.
\end{aligned}$$

The Jacobi Symbol and Primality Test[a]

A result generalizing Proposition 10.3 in the book:

Theorem 62 *The group of set $\Phi(n)$ under multiplication* mod n *has a primitive root if and only if n is either 1, 2, 4, p^k, or $2p^k$ for some nonnegative integer k and and odd prime p.*

This result is essential in the proof of the next lemma.

Lemma 63 *If $(M|N) = M^{(N-1)/2}$* mod N *for all $M \in \Phi(N)$, then N is prime. (Assume N is odd.)*

[a]Clement Hsiao (R88067) pointed out that the textbook's proof in Lemma 11.8 is incorrect while he was a senior in January 1999.

The Number of Witnesses to Compositeness

Theorem 64 (Solovay and Strassen (1977)) *If N is an odd composite, then $(M|N) \neq M^{(N-1)/2}$* mod N *for at least half of $M \in \Phi(N)$.*

- By Lemma 63 there is at least one $a \in \Phi(N)$ such that $(a|N) \neq a^{(N-1)/2} \bmod N$.
- Let $B = \{b_1, b_2, \ldots, b_k\} \subseteq \Phi(N)$ be the set of all distinct residues such that $(b_i|N) = b_i^{(N-1)/2} \bmod N$.
- Let $aB = \{ab_i \bmod N : i = 1, 2, \ldots, k\}$.

The Proof (concluded)

- $|aB| = k$.
  - $ab_i = ab_j \bmod N$ implies $N | a(b_i - b_j)$, which is impossible because $\gcd(a, N) = 1$ and $N > |b_i - b_j|$.
- $aB \cap B = \emptyset$ because

$$(ab_i)^{(N-1)/2} = a^{(N-1)/2} b_i^{(N-1)/2} \neq (a|N)(b_i|N) = (ab_i|N).$$

- Combining the above two results, we know

$$\frac{|B|}{\phi(N)} \leq 0.5.$$

if N is even but $N \neq 2$ **then**
  return "N is a composite";
else if $N = 2$ **then**
  return "N is a prime";
end if
Pick $M \in \{2, 3, \ldots, N-1\}$ randomly;
if $\gcd(M, N) > 1$ **then**
  return "N is a composite";
else
  if $(M|N) \neq M^{(N-1)/2} \bmod N$ **then**
    return "N is a composite";
  else
    return "N is a prime";
  end if
end if

Analysis

- The algorithm certainly runs in polynomial time.
- There are no false positives (for COMPOSITENESS).
  - When the algorithm says the number is a composite, it is always correct.
- The probability of a false negative is at most one half.
  - When the algorithm says the number is a prime, it may err.
  - If the input is a composite, then the probability that the algorithm errs is one half.
- The error probability can be reduced but not eliminated.

The Improved Density Attack for COMPOSITENESS