Monte Carlo Algorithms

- 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 negatives**).
- The probability that the algorithm makes a false negative is at most 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.

The Markov Inequality[a]

Lemma 53 *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 $k \times T(n)$ 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 correct answer with probability at least $1 - 1/k$.
- By running this algorithm m times, we reduce the error probability to $\leq k^{-m}$.

A Random Walk Algorithm for ϕ in CNF Form

Start with an *arbitrary* truth assignment T;
for $i = 1, 2, \ldots, r$ **do**
  if $T \models \phi$ **then**
    return "ϕ is satisfiable";
  else
    Let c be an unsatisfiable clause in ϕ under T; {All of its literals are false under T.}
    Pick any x of these literals *at random*;
    Modify T to make x true;
  end if
end for
return "ϕ is probably unsatisfiable";

3SAT and 2SAT Again

- Note that if ϕ is unsatisfiable, the algorithm will not refute it.
- The random walk algorithm runs in exponential time for 3SAT.
- But we will show that it works well for 2SAT.

Theorem 54 *Suppose that the random walk algorithm with* $r = 2n^2$ *is applied to any satisfiable* 2SAT *problem with* n *variables. Then a satisfying truth assignment will be discovered with probability at least 0.5.*

The Proof

- Let $\hat{T}$ be a truth assignment such that $\hat{T} \models \phi$.
- Let $t(i)$ denote the expected number of repetitions of the flipping step until a satisfying truth assignment is found if our starting T differs from $\hat{T}$ in i values.
  - Their Hamming distance is i.
- It can be shown that $t(i)$ is finite.
- $t(0) = 0$ because it means that $T = \hat{T}$ and hence $T \models \phi$.
- If $T \neq \hat{T}$ or T is not equal to any other satisfying truth assignment, then we need to flip at least once.

The Proof (continued)

- We flip to pick among the 2 literals of a clause not satisfied by the present T.
- At least one of the 2 literals is true under $\hat{T}$, because $\hat{T}$ satisfies all clauses.
- So we have at least 0.5 chance of moving closer to $\hat{T}$.
- Thus

$$t(i) \leq \frac{t(i-1) + t(i+1)}{2} + 1$$

  for $0 < i < n$.
- Inequality is used because, for example, T may differ from $\hat{T}$ in both literals.

The Proof (continued)

- It must also hold that

$$t(n) \leq t(n-1) + 1$$

  because at $i = n$, we can only decrease i.

- As we are only interested in upper bounds, we solve

$$\begin{aligned} x(0) &= 0 \\ x(n) &= x(n-1) + 1 \\ x(i) &= \frac{x(i-1) + x(i+1)}{2} + 1, \quad 0 < i < n \end{aligned}$$

- This is one-dimensional random walk with a reflecting and an absorbing barrier.

The Proof (continued)

- Add the equations up to obtain

$$
\begin{aligned}
& x(1) + x(2) + \cdots + x(n) \\
= \quad & \frac{x(0) + x(1) + 2x(2) + \cdots + 2x(n-2) + x(n-1) + x(n)}{2} \\
& +n + x(n-1).
\end{aligned}
$$

- Simplify to yield

$$\frac{x(1) + x(n) - x(n-1)}{2} = n.$$

- As $x(n) - x(n-1) = 1$, we have

$$x(1) = 2n - 1.$$

The Proof (continued)

- Iteratively, we obtain

$$
\begin{aligned}
x(2) &= 4n - 4 \\
&\vdots \\
x(i) &= 2in - i^2
\end{aligned}
$$

- The worst case happens when $i = n$, in which case

$$x(n) = n^2.$$

The Proof (continued)

- We therefore reach the conclusion that

$$t(i) \leq x(i) \leq x(n) = n^2.$$

- So the expected number of steps is at most n^2.
- The algorithm picks a running time $2n^2$.
- This amounts to invoking the Markov inequality (p. 282) with $k = 2$, with the consequence of having 0.5 probability.

Boosting the Performance

- We can pick $r = 2mn^2$ to have an error probability of $\leq (2m)^{-1}$ by Markov's inequality.
- Alternatively, with the same running time, we can run the $r = 2n^2$ algorithm m times.
- But the error probability is reduced to $\leq 2^{-m}$!
- The gain comes from the fact that Markov's inequality does not take advantage of any specific feature of the random variable.
- The gain also comes from the fact that the two algorithms are different.

The Fermat Test

- Fermat's "little" theorem on p. 262 suggests the following primality test for any given number p:
  - Pick a number a randomly from $\{1, 2, \ldots, p-1\}$.
  - If $a^{p-1} \neq 1 \bmod p$, then declare "p is composite."
  - Otherwise, declare "p is probably prime."
- Unfortunately, there are composite numbers called **Carmichael numbers** that will pass the Fermat test for *all* $a \in \{1, 2, \ldots, p-1\}$.
- It is only recently that Carmichael numbers are known to be infinite in number.

Square Roots Modulo a Prime

- Equation $x^2 = a \bmod p$ has at most two (distinct) roots by Lemma 50 on p. 264.
  - The roots are called **square roots**.
  - Numbers a with square roots and $\gcd(a, p) = 1$ are called **quadratic residues**:
    $1^2 \bmod p, 2^2 \bmod p, \ldots, (p-1)^2 \bmod p$.
  - $x^2 = a \bmod p$ has at most two roots when p is odd.
- We shall show that a number either has two roots or has none, and testing which is true is trivial.

Euler's Test

Lemma 55 (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.
- If $a = r^{2j}$, then $a^{(p-1)/2} = r^{j(p-1)} = 1 \bmod p$ and its two distinct roots are $r^j, -r^j (= r^{j+(p-1)/2})$.

The Proof (continued)

- 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 are taken.
- 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, $r^{(p-1)/2} = -1 \bmod p$.
- $a^{(p-1)/2} = (r^{(p-1)/2})^{2j+1} = (-1)^{2j+1} = -1 \bmod p$.

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

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

- Euler's test implies $a^{(p-1)/2} \equiv (a \,|\, p) \bmod p$ 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 56 (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 $p|(j-i)$ or $p|q$.
- No two elements of R add up to p.
  - If $iq + jq = 0 \bmod p$, then $p|(i+j)$ 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 (continued)

- 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, $R' = \{\pm iq : 1 \leq i \leq (p-1)/2\}$, where exactly m of the elements have the minus sign.
- Taking the product of all elements in the two representations of R', we have $[(p-1)/2]! = (-1)^m q^{(p-1)/2}[(p-1)/2]! \bmod p$.
- Because $\gcd([(p-1)/2]!, p) = 1$, the lemma follows.

Legendre's Law of Quadratic Reciprocity

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

Lemma 57 (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

$$\sum_{i=1}^{(p-1)/2} i = \frac{(p-1)(p+1)}{8} \bmod 2.$$

The Proof (continued)

- On the other hand, the sum equals

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

  - Signs are irrelevant under mod2.

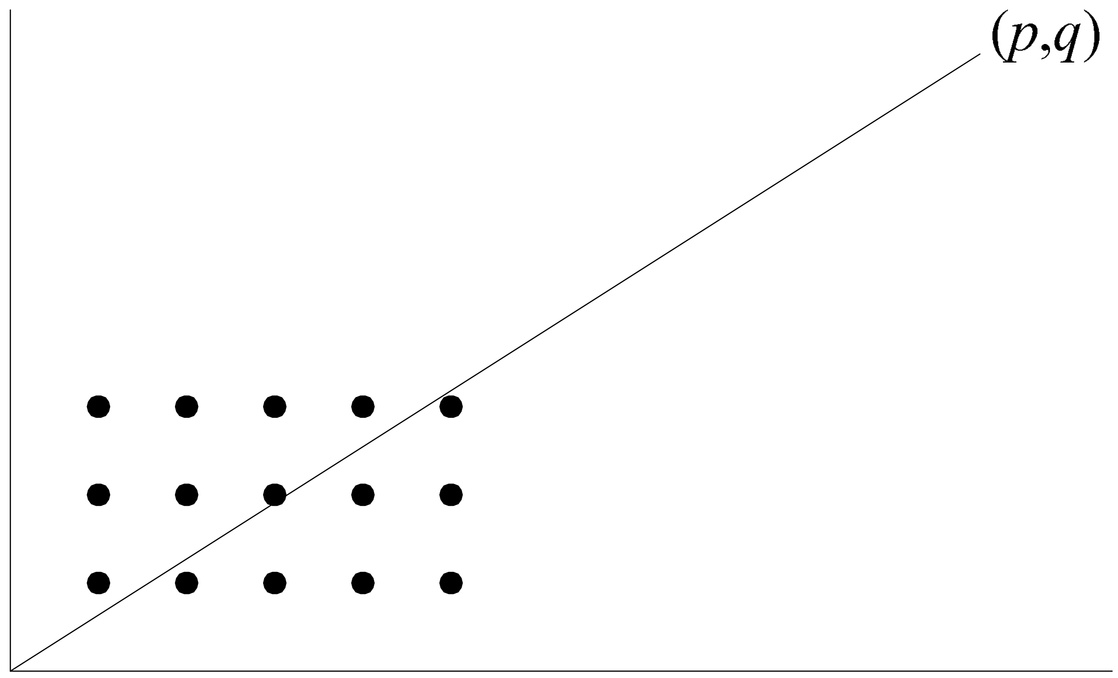
- After ignoring odd multipliers and noting that the first term above equals $\sum_{i=1}^{(p-1)/2} i$:

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

The Proof (continued)

- $m = \sum_{i=1}^{(p-1)/2} \lfloor \frac{iq}{p} \rfloor$ 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).
- From Gauss's lemma on p. 298, $(q|p)$ 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 $\frac{p-1}{2} \cdot \frac{q-1}{2}$.

Eisenstein's Rectangle



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

The Jacobi Symbol[a]

- The Legendre symbol only works for an odd prime modulus.
- The **Jacobi symbol** $(a \mid m)$ 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 is odd and is greater than one, then

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

- Define $(a \mid 1) = 1$.

[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. $(ab \,|\, m) = (a \,|\, m)(b \,|\, m)$.
2. $(a \,|\, m_1 m_2) = (a \,|\, m_1)(a \,|\, m_2)$.
3. If $a \equiv b \bmod m$, then $(a \,|\, m) = (b \,|\, m)$.
4. $(-1 \,|\, m) = (-1)^{(m-1)/2}$.
5. $(2 \,|\, m) = (-1)^{(m^2-1)/8}$.
6. If a and m are both odd, then $(a \,|\, m)(m \,|\, a) = (-1)^{(a-1)(m-1)/4}$.

Calculation of $(2200|999)$

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

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

The Jacobi Symbol and Primality Test[a]

Lemma 58 *If* $(M|N) = M^{(N-1)/2} \bmod N$ *for all* $M \in \Phi(N)$*, then* N *is prime. (Assume* N *is odd.)*

- First assume that $N = rp^a$, where p is an odd prime, $\gcd(r, p) = 1$, $r > 1$ (not necessarily prime), and a is odd.
- We shall derive a contradiction.
- By the assumption,

$$M^{(N-1)/2} = \pm 1 \bmod N \text{ for all } M \in \Phi(N). \tag{3}$$

- Suppose $M^{(N-1)/2} = -1 \bmod N$ for some $M \in \Phi(N)$.

[a]Clement Hsiao pointed out that the textbook's proof in Lemma 11.8 is incorrect when he was a senior.

The Proof (continued)

- Then there is a unique M' such that

$$\begin{aligned} M' &= 1 \bmod r \\ M' &= M \bmod p^a \end{aligned}$$

by the Chinese remainder theorem.

- As $\gcd(1, r) = \gcd(M, p^a) = 1$, we have $M' \in \Phi(N)$.

- Now

$$\begin{aligned} {M'}^{(N-1)/2} &= 1 \bmod r \\ {M'}^{(N-1)/2} &= -1 \bmod p^a \end{aligned}$$

The Proof (continued)

- But $M'^{(N-1)/2} \neq \pm 1 \bmod N$.
  - Otherwise,

$$M'^{(N-1)/2} \bmod r = M'^{(N-1)/2} \bmod p^a.$$

- This contradicts Eq. (3).
- Hence

$$M^{(N-1)/2} = 1 \bmod N \quad \text{for all } M \in \Phi(N). \tag{4}$$

The Proof (continued)

- By the Chinese remainder theorem again, there is a unique $M' \in \{0, 1, \ldots, rp-1\}$ such that

$$\begin{aligned} M' &= 1 \bmod r \\ M' &= z \bmod p \end{aligned}$$

  where z is one of the quadratic nonresidues modulo p.

- As $\gcd(1, r) = \gcd(z, p) = 1$, $M' \in \Phi(rp)$ and so $M' \in \Phi(N)$.

- $(M'|N) = (M'|r)(M'|p^a) = (M'|r)(M'|p)^a = (1|r)(z|p) = -1$, contradiction Eq. (4).

The Proof (continued)

- Second, assume $N = p_1 p_2 \cdots p_k$, where p_i are distinct odd primes.
- Let $r \in \Phi(p_1)$ such that $(r \mid p_1) = -1$.
- By the Chinese remainder theorem, there is an $M \in \Phi(N)$ such that

$$\begin{aligned} M &= r \bmod p_1 \\ M &= 1 \bmod p_i, \quad 2 \leq i \leq k \end{aligned}$$

The Proof (continued)

- By the hypothesis,

$$M^{(N-1)/2} = (M \mid N) = \prod_{i=1}^{k} (M \mid p_i) = -1 \bmod N.$$

- Hence

$$M^{(N-1)/2} = -1 \bmod p_2.$$

- But because $M = 1 \bmod p_2$,

$$M^{(N-1)/2} = 1 \bmod p_2,$$

  a contradiction again.

The Number of Witnesses to Compositeness

Theorem 59 *If N is an odd composite, then* $(M|N) \neq M^{(N-1)/2} \bmod N$ *for at least half of $M \in \Phi(N)$.*

- By Lemma 58 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 (continued)

- $|aB| = k$.
  - $ab_i = ab_j \bmod N$ 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 $|B|/\phi(N) \leq 0.5$.

A Polynomial-Time Randomized Algorithm for Primality (or Compositeness)[a]

```
Pick M ∈ {2, 3, ..., N − 1} randomly;
if gcd(M, N) > 1 then
    return "N is a composite";
else
    if (M|N) ≠ M^((N−1)/2) mod N then
        return "N is a composite";
    else
        return "N is probably a prime";
    end if
end if
```

[a]Solovay, Strassen, 1977.

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 probability of error can be reduced but not eliminated.