Theory of Computation
Chapter 11: Randomized Computation

Guan-Shieng Huang

Dec. 6, 2006

Outline

- Basic Concept
- Examples
- Complexity Classes
- Basic Techniques

Randomized Computation

1. Can random numbers help us solve computational problems?
2. In a randomized algorithm, we may make the following statement:
   (a) Given any number $n > 2$, we can decide whether n is prime with high probability.

Types of Errors

- positive: when answer "yes"
  negative: when answer "no"
- true positive; true negative:
  The answer coincides with the fact
- false positive; false negative
  The answer is wrong

Example

1. Given $n = 5$, suppose we want to decide whether $n > 4$.
   If we answer "no", then this answer is a false negative;
   if we answer "yes", then this answer is a true positive.
2. Suppose we want to decide whether n is even.
   Answer "yes" $\Longrightarrow$ false positive; answer "no" $\Longrightarrow$ true negative.

Monte Carlo Algorithm

A randomized algorithm that never appears false positive.

- If it answers “yes”, the answer must be correct.
- If it answers “no”, the answer may be wrong.
- With high probability that it can answer “yes” if it is really this case.

Remark Monte Carlo method or Monte Carlo simulation is a rather general term referring to a procedure that involves randomness.

Examples

- Symbolic Determinants
- Random Walks for 2SAT
- Compositeness

Symbolic Determinants

- Let A be an $n \times n$ matrix with each entry a multi-variate polynomial. $(x^3y + 3y^5z)$
  We want to determine whether the determinant of A is not a zero polynomial.

- $\det A = \sum_\pi \sigma(\pi) \prod_{i=1}^n a_{i,\pi(i)}$ where $A = (a_{i,j})_{n \times n}$; $\sigma(\pi) = 1$ if π is an even permutation, -1 if π is odd.

$$\det A = \sum_{\pi} \sigma(\pi) \prod_{i=1}^{n} a_{i,\pi(i)}$$

$$\begin{aligned} \det & \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \\ = & \quad a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} \\ & -a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1} \end{aligned}$$

- $\pi = [3, 2, 1]$ is an odd permutation.
  $a_{1,\pi(1)}a_{2,\pi(2)}a_{3,\pi(3)} = a_{1,3}a_{2,2}a_{3,1}$
- $\pi = [2, 3, 1]$ is an even permutation.
  $a_{1,\pi(1)}a_{2,\pi(2)}a_{3,\pi(3)} = a_{1,2}a_{2,3}a_{3,1}$

- Gaussian elimination can solve “numerical determinants” in polynomial time.

- No body knows how to solve the symbolic determinants in polynomial time.

Randomized Algorithm for Symbolic Determinants

Assume there are m variables in A and the highest degree if each variable in the expansion is at most d.

1. Choose m random integers $i_1, \ldots, i_m$ between 0 and $M = 2md$.
2. Compute the determinant $\det A(i_1, \ldots, i_m)$ by Gaussian elimination.
3. If the result$\neq 0$, reply "yes".
4. If the result$= 0$, reply "probably equal to 0".

Lemma 11.1 Let $p(x_1, \ldots, x_m)$ be a polynomial, not identically zero, in m variables each of degree at most d in it, and let $M > 0$ be an integer. Then the number of m-tuples $(x_1, \ldots, x_m) \in \mathbb{Z}_M^m$ such that $p(x_1, \ldots, x_m) = 0$ is at most mdM^{m-1}.

Proof.

1. By induction on m. When $m = 1$ the lemma says that no polynomial of degree $\leq d$ can have more than d roots.
2. Suppose the result is true for $m - 1$ variables.
   Let the degree of x_m is $t \leq d$. We can rewrite $p(x_1, \ldots, x_m)$ as $q(x_1, \ldots, x_{m-1})x_m^t + r(x_1, \ldots, x_m)$. Consider $x_1, \ldots, x_{m-1}$ according to whether they can make $q(x_1, \ldots, x_{m-1}) = 0$.

$$\#\text{roots} \leq (m-1)dM^{m-2} \cdot M + M^{m-1}t \leq mdM^{m-1}.$$

Random Walks for 2SAT

2SAT: Satisfiability problem with each clause containing at most two literals.

Algorithm

1. Start with any truth assignment T.
2. Repeat the following steps r times.
   (a) If there is no unsatisfied clause, reply "Formula is satisfiable" and halt.
   Otherwise, pick any unsatisfied clause, flip the value of any one literal inside it.
3. Reply "Formula is probably unsatisfiable".

Theorem Let $r = 2n^2$. Then this algorithm can find a satisfiable truth assignment with probability at least $\frac{1}{2}$ when the 2SAT formula is satisfiable.

Proof.

1. $\widehat{T}$: a satisfying truth assignment for this formula
   T: current assignment
2. $t(i)$: the expectation for the number of flipping if T differs from $\widehat{T}$ in exactly i values
3. $t(0) = 0$
   $t(i) \leq \frac{1}{2}(t(i-1) + t(i+1)) + 1$
   $t(n) = t(n-1) + 1$

4. Let $x(0) = 0$ $x(i) = \frac{1}{2}(x(i-1) + x(i+1)) + 1$
   $x(n) = x(n-1) + 1$
   Then $t(i) \leq x(i) = 2in - i^2 \leq n^2$.

5. Let $r = 2n^2$. Then $\texttt{Prob}[r \geq 2n^2] \leq \frac{1}{2}$.

Lemma 11.2 (Markov Inequality) If x is a non-negative random variable, then for any $k > 0$, $\texttt{Prob}[x \geq k\mu_x] \leq \frac{1}{k}$ where μ_x is the expectation of x.

Proof. (discrete case)

$$\mu_x = \sum_i ip_i = \sum_{i<k\mu_x} ip_i + \sum_{i\geq k\mu_x} ip_i \geq k\mu_x \texttt{Prob}[x \geq k\mu_x].$$

$$\therefore \texttt{Prob}[x \geq k\mu_x] \leq \frac{1}{k}.$$

Fermat Test

1. If n is prime, then $a^{n-1} \equiv 1 \pmod{n}$ for all a not divided by n.
2. Hypothesis: n is not prime $\Longrightarrow$ at least half of nonzero residues a can make $a^{n-1} \not\equiv 1 \pmod{n}$.
3. If it is true, we would have a polynomial Monte Carlo algorithm for testing whether n is composite.
   Unfortunately, this statement is false.

Square Roots Modulo a Prime

Consider $x^2 \equiv a \pmod{p}$ where $p \geq 3$. Then exactly half of the nonzero residues have square roots.

Proof.

- Consider the squares of $1, 2, 3, \ldots, p-1$.
- They are exactly those numbers that have square roots.
- k and $p-k$ collapse after squaring.
- However, $x^2 \equiv a$ has at most two roots, and in fact, either zero or two distinct roots.

Lemma 11.3 If $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$, then $x^2 \equiv a$ has two roots. Otherwise, $a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$ and it has no roots.

Proof. Let r be a primitive root for p. Then each nonzero residue $a \equiv r^k$ for some $k \geq 0$.

1. $k = 2j$: $a^{\frac{p-1}{2}} \equiv (r^{2j})^{\frac{p-1}{2}} = (r^{p-1})^j \equiv 1$, and the square roots for a are r^j and $r^{j+\frac{p-1}{2}}$.
2. $k = 2j+1$: $a^{\frac{p-1}{2}} = (r^{2j+1})^{\frac{p-1}{2}} = r^{j(p-1)+\frac{p-1}{2}} \equiv r^{\frac{p-1}{2}} \equiv -1 \pmod{p}$, and it has no square roots.

Legendre Symbol

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ has square roots in } p \\ 0 & \text{if } p \text{ divides } a \\ -1 & \text{if } a \text{ has no seuqre root in } p \end{cases}$$

for prime numbers $p > 2$.

Theorem $\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \mod p$.

Corollary $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Gauss's Lemma

$\left(\frac{a}{p}\right) = (-1)^m$ where $m = |\{i: \ 1 \le i \le \frac{p-1}{2}, qi \mod p > \frac{p-1}{2}\}|$ and $p > 2$.

Proof.

Consider

$$q, 2q, 3q, \ldots, \frac{p-1}{2} \cdot q$$

and

$$-\frac{p-1}{2}, \ldots, -1, 0, 1, \ldots, \frac{p-1}{2}.$$

Either k or $-k$ $(1 \le k \le \frac{p-1}{2})$ can be mapped by one number qi, but not both:

$$qi \equiv -qj \pmod{p} \Rightarrow q(i+j) \equiv 0 \pmod{p} \Rightarrow p|(i+j).$$

And no two numbers qi and qj can be the same:

$$qi \equiv qj \pmod{p} \Rightarrow p|i-j.$$

$$\prod_{1\leq i\leq \frac{p-1}{2}} qi = (\frac{p-1}{2})! \cdot q^{\frac{p-1}{2}} \equiv (-1)^m (\frac{p-1}{2})!$$

$$\therefore (-1)^m \equiv q^{\frac{p-1}{2}} \equiv \left(\frac{q}{p}\right) \pmod{p}.$$

Legendre's Law of Reciprocity

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \text{ if } \gcd(p,q) = 1.$$

Proof.

1.

$$1 + 2 + 3 + \cdots + \frac{p-1}{2} \equiv \sum_{i=1}^{\frac{p-1}{2}} (qi - p\left\lfloor \frac{qi}{p} \right\rfloor) + mp \pmod 2.$$

$$\because 0 \le a \le \tfrac{p-1}{2} \Rightarrow p - a = a + p - 2a \equiv a + p \equiv a + 1 \pmod 2.$$

2.

$$\therefore \sum_{i=1}^{\frac{p-1}{2}} i \equiv q \sum_{i=1}^{\frac{p-1}{2}} i - p \sum_{i=1} \frac{p-1}{2} \left\lfloor \frac{qi}{p} \right\rfloor + mp \pmod 2$$

3.

$$\because p \equiv q \equiv 1 \pmod 2,$$

$$\therefore m \equiv \sum_{i=1}^{\frac{p-1}{2}} \left\lfloor \frac{qi}{p} \right\rfloor \pmod 2$$

4. No grid lies inside $(0,0)$—(p,q). Hence,

$$m + m' \equiv \sum_{i=1}^{\frac{p-1}{2}} \left\lfloor \frac{qi}{p} \right\rfloor + \sum_{j=1}^{\frac{q-1}{2}} \left\lfloor \frac{pj}{q} \right\rfloor \equiv \frac{p-1}{2} \cdot \frac{q-1}{2} \pmod 2.$$

5.

$$\therefore \left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^m \cdot (-1)^{m'} = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Jacob's Symbol

$$\left(\frac{M}{N}\right) = \left(\frac{M}{p_1}\right)\left(\frac{M}{p_2}\right)\cdots\left(\frac{M}{p_n}\right)$$

if $N = p_1 p_2 \ldots p_n$ where p_i's are odd primes (which may be the same).

Lemma 11.6

1. $\left(\frac{M_1 M_2}{N}\right) = \left(\frac{M_1}{N}\right)\left(\frac{M_2}{N}\right)$
2. $\left(\frac{M+N}{N}\right) = \left(\frac{M}{N}\right)$
3. $\left(\frac{N}{M}\right)\left(\frac{M}{N}\right) = (-1)^{\frac{M-1}{2}\frac{N-1}{2}}$ if $\gcd(M, N) = 1$ and M, N are odd.

Proof.

1. $\left(\frac{M_1 M_2}{N}\right) = \prod_i \left(\frac{M_1 M_2}{p_i}\right) = \prod_i \left(\frac{M_1}{p_1}\right) \prod_j \left(\frac{M_2}{p_j}\right) = \left(\frac{M_1}{N}\right)\left(\frac{M_2}{N}\right)$

2. $\left(\frac{M+N}{N}\right) = \prod_i \left(\frac{M+N}{p_i}\right) = \prod M p_i = \left(\frac{M}{N}\right)$

3. $\left(\frac{M}{N}\right)\left(\frac{N}{M}\right) = \prod_{i,j}\left(\frac{q_j}{p_i}\right) \cdot \prod_{i,j}\left(\frac{p_i}{q_j}\right) = \prod_{i,j}\left[\left(\frac{q_j}{p_i}\right)\left(\frac{p_i}{q_j}\right)\right]$

$$= \prod_{i,j}(-1)^{\frac{p_i-1}{2}\cdot\frac{q_j-1}{2}} = (-1)^{\sum_{i,j}\frac{p_i-1}{2}\frac{q_j-1}{2}}.$$

And $\sum_{i,j}\frac{p_i-1}{2}\frac{q_j-1}{2} = \sum_i \frac{p_i-1}{2}\sum_j \frac{q_j-1}{2}$, and $\frac{a-1}{2} + \frac{b-1}{2} \equiv \frac{ab-1}{2} \pmod 2$.

$$\therefore \sum_j \frac{q_j-1}{2} \equiv \frac{M-1}{2} \pmod 2,$$

$$\text{and} \sum_i \frac{p_i-1}{2} \equiv \frac{N-1}{2} \pmod 2.$$

Lemma

$$\left(\frac{2}{M}\right) = (-1)^{\frac{M^2-1}{8}}$$

Proof.

Let $M = q_1 \dots q_m$. We first show that $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$ for odd primes p.

Consider $2, 2 \times 2, \dots, 2i, \dots, 2 \times \frac{p-1}{2}$ for $1 \le i \le i \le \frac{p-1}{2}$.

$2i > \frac{p-1}{2} \Rightarrow i > \frac{p-1}{4}$

$$\therefore m = \frac{p-1}{2} - \left\lfloor \frac{p-1}{4} \right\rfloor = \frac{p-1}{2} + \left\lceil -\frac{p-1}{4} \right\rceil$$

$$= \left\lceil \frac{p-1}{2} - \frac{p-1}{4} \right\rceil = \left\lceil \frac{p-1}{4} \right\rceil \equiv \frac{p^2-1}{8} \pmod{2}.$$

Lemma Given two integers M and N with $\ell = \lg MN$, $\gcd(M, N)$ and $\left(\frac{M}{N}\right)$ can be computed in $O(\ell^3)$ time.

Summary

1. $\left(\frac{M}{N}\right) = 0$ if $\gcd(M, N) \neq 1$;
2. $\left(\frac{M_1 M_2}{N}\right) = \left(\frac{M_1}{N}\right)\left(\frac{M_2}{N}\right)$; $\left(\frac{M^2}{N}\right) = 1$;
3. $\left(\frac{M}{N}\right) = -\left(\frac{N}{M}\right)$ iff $M \equiv N \equiv 3 \pmod 4$; $\left(\frac{M}{N}\right) = \left(\frac{N}{M}\right)$ otherwise;
4. $\left(\frac{2}{N}\right) = -1$ iff $N \equiv 3 \pmod 8$ or $N \equiv 5 \pmod 8$.

Example

$$\left(\frac{163}{511}\right) = -\left(\frac{511}{163}\right) = -\left(\frac{22}{163}\right) = -\left(\frac{2}{163}\right)\left(\frac{11}{163}\right)$$

$$= \left(\frac{11}{163}\right) = -\left(\frac{163}{11}\right) = -\left(\frac{9}{11}\right) = -\left(\frac{11}{9}\right) = -\left(\frac{2}{9}\right) = -1.$$

Lemma 11.8 If $\left(\frac{M}{N}\right) \equiv M^{\frac{N-1}{2}} \pmod{N}$ for all $M \in \Phi(N)$, then N is prime.

Proof.

Suppose N is composite.

1. $N = p_1 p_2 \ldots p_k$, the product of distinct primes.
   Let r be a number such that $\left(\frac{r}{p_1}\right) = -1$,
   $r \mod p_j = 1$ for $2 \leq j \leq k$.
   Then $r^{\frac{N-1}{2}} \equiv \left(\frac{r}{N}\right) \equiv \prod \left(\frac{r}{p_i}\right) = -1 \pmod{N}$.
   Hence $r^{\frac{N-1}{2}} \equiv 1 \pmod{p_2}$, but $r^{\frac{N-1}{2}} \equiv 1^{\frac{N-1}{2}} \equiv 1 \pmod{p_2}$, contradiction.

2. Let $N = p^2 m$ for some $p > 2$ and $m > 1$.
   Let r be a primitive root for p^2. Then $\phi(p^2) = p(p-1) | N - 1$.
   Hence $p | N - 1$ and $p | N$, absurd.

Lemma 11.2 If N is an odd composite, then for at least half of $M \in \Phi(N)$, $\left(\frac{M}{N}\right) \not\equiv M^{\frac{N-1}{2}} \pmod N$.

Proof.

By Lemma 11.8, there is at least one $a \in \Phi(N)$ such that

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

Let $B \subseteq \Phi(N)$ such that $\left(\frac{b}{N}\right) \equiv b^{\frac{N-1}{2}} \pmod N$ for all $b \in B$.

Let $a \cdot B$ be $\{ab : \ b \in B\}$.

Then $(ab)^{\frac{N-1}{2}} \equiv a^{\frac{N-1}{2}} \cdot b^{\frac{N-1}{2}} \not\equiv \left(\frac{a}{N}\right)\left(\frac{b}{N}\right) = \left(\frac{ab}{N}\right) \pmod N$.

The size of B and aB are the same.

Hence at least half of $M \in \Phi(N)$ make $\left(\frac{M}{N}\right) \not\equiv M^{\frac{N-1}{2}} \pmod N$.

Monte Carlo Algorithm for Compositeness

Algorithm

Input N:

1. If $2|N$, reply "Composite".
2. Generate a random number M between 2 and $N-1$.
   If $\gcd(M, N) \neq 1$, reply "Composite".
3. If $\left(\frac{M}{N}\right) \not\equiv M^{\frac{N-1}{2}}$, "Composite".
4. Reply "Probably prime".

This algorithm takes cubic time.

Randomized Complexity Classes

- RP: Randomized Polynomial time
- ZPP: Zero-error Probabilistic Polynomial time
- BPP: Bounded Probabilistic Polynomial time

RP

(Randomized Polynomial time)

Modelled as a non-deterministic Turing machine with

1. each computation on an input of length n terminates at $p(n)$ steps;
2. if $x \in L$, then at least half of the computations halts with "yes";
3. if $x \notin L$, then all computations halts with "no".

Remark Condition 2 can be relaxed to $\Omega(\frac{1}{p(n)})$.
Suppose the probability of false negative is at most $1-\eta$.

- Repeating the RP algorithm k times can reduce the probability $\leq (1-\eta)^k$.
- Let $k = \lceil \log_{(1-\eta)} \frac{1}{2} \rceil = \lceil -\frac{1}{\lg(1-\eta)} \rceil$, which makes $(1-\eta)^k \leq \frac{1}{2}$.
- $\lg(1-\eta) \approx -\frac{\eta}{\ln 2}$,
  $\therefore k \approx -\frac{1}{\lg(1-\eta)} \approx \frac{\ln 2}{\eta} = O(p(n))$ when $\eta = \Omega(\frac{1}{p(n)})$.

ZPP

(Zero-error Probabilistic Polynomial time = RP$\cap$coRP)

It meas that there are two RP algorithms, one for $x \in L$ and the other for $x \in \bar{L}$.

BPP

(Bounded Probabilistic Polynomial time)

$$\begin{cases} \texttt{Prob}R(x) = \text{“yes”} \geq \frac{3}{4} & \text{if } x \in L \\ \texttt{Prob}R(x) = \text{“no”} \geq \frac{3}{4} & \text{if } x \notin L \end{cases}$$

Remark The condition can be relaxed to

$$\begin{cases} \texttt{Prob}R(x) = \text{“yes”} \geq \frac{1}{2} + \epsilon & \text{if } x \in L \\ \texttt{Prob}R(x) = \text{“no”} \geq \frac{1}{2} + \epsilon & \text{if } x \notin L \end{cases}$$

where $\epsilon = \Omega(\frac{1}{p(n)})$.

The Chernoff Bound

(Estimate the tail probability of independent Bernoulli trials.)

- $x_1, \ldots, x_n$: independent random variables taking values 1 and 0 with prob. p and $1-p$, respectively.
- $X = \sum_{i=1}^{n} x_i$
- $0 \leq \theta \leq 1$

then $\texttt{Prob} X \geq (1+\theta)pn \leq \exp(-\frac{\theta^2}{3}pn)$.

Corollary Let $p = \frac{1}{2} + \epsilon$ for some $\epsilon > 0$.
Then $\texttt{Prob} \sum_{i=1}^{n} x_i \leq \frac{n}{2} \leq \exp(-\frac{\epsilon^2 n}{6})$.

Random Sources

Do we have true random sources?

- Pseudo randomness
- Perfect random source
- Slightly random source

Derandomization

Make a randomized algorithm deterministic without losing much efficiency.