CS 473: Algorithms

Chandra Chekuri
chekuri@cs.illinois.edu
3228 Siebel Center

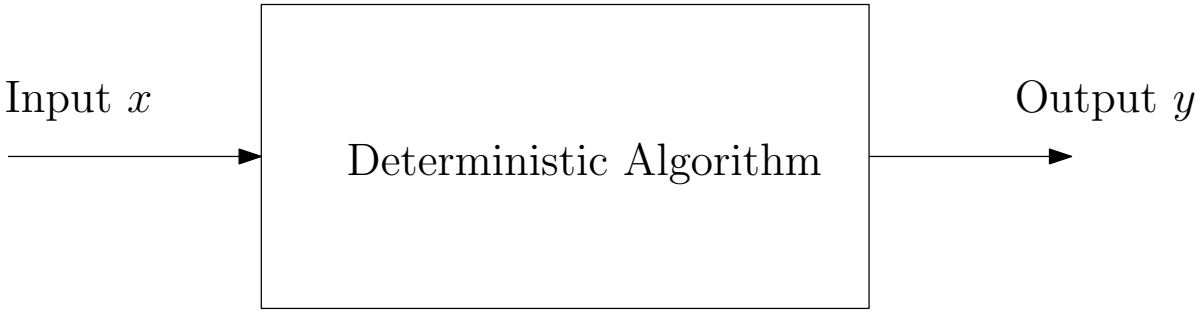
University of Illinois, Urbana-Champaign

Fall 2010

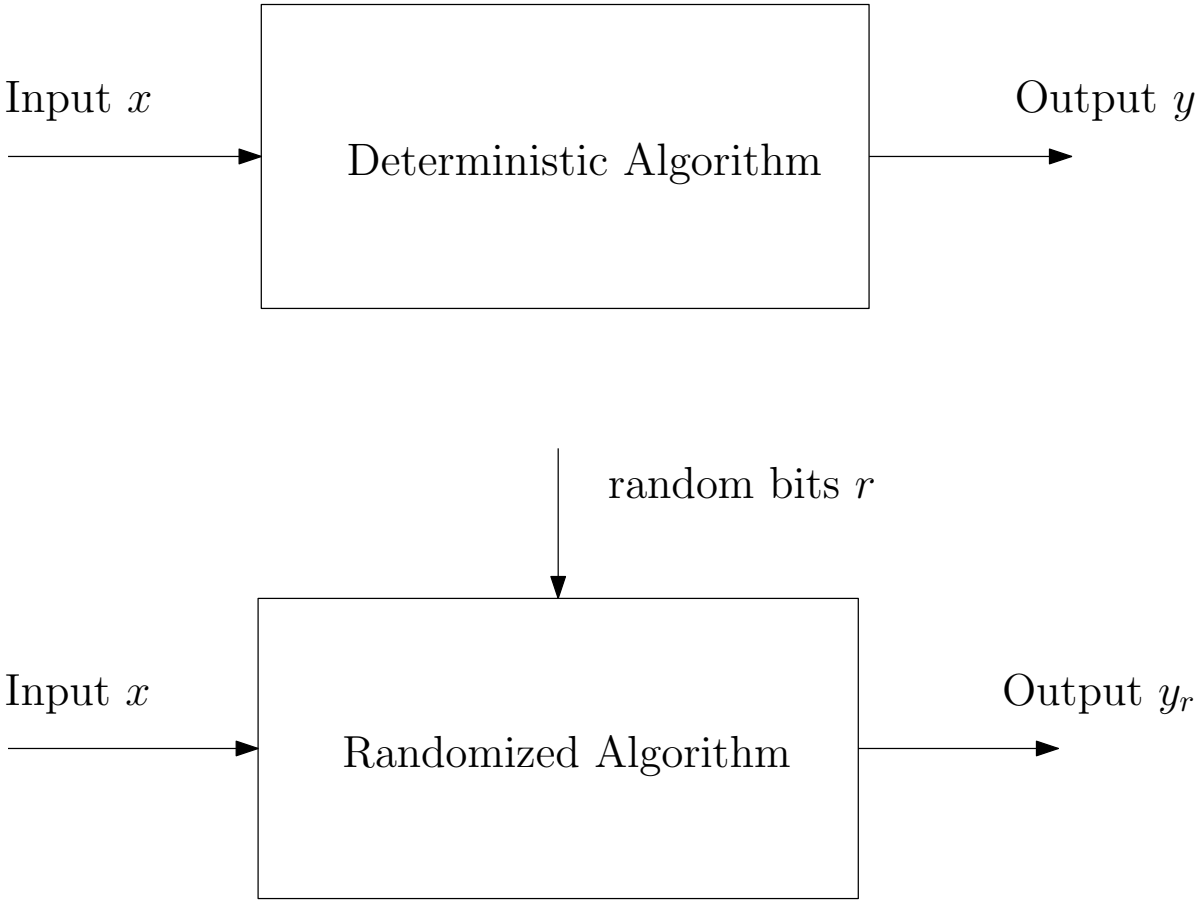
Part I

Introduction to Randomized Algorithms

Randomized Algorithms



Randomized Algorithms



Example: Randomized Quicksort

Quick Sort[Hoare]

1. Pick a pivot element from array
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself.
3. Recursively sort the subarrays, and concatenate them.

Randomized Quick Sort

1. Pick a pivot element *uniformly at random* from the array
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself.
3. Recursively sort the subarrays, and concatenate them.

Example: Randomized Quicksort

Recall: Quick Sort can take $\Omega(n^2)$ time to sort array of size n.

Example: Randomized Quicksort

Recall: Quick Sort can take $\Omega(n^2)$ time to sort array of size n.

Theorem

Randomized Quick Sort sorts a given array of length n in $O(n \log n)$ expected time.

Example: Randomized Quicksort

Recall: Quick Sort can take $\Omega(n^2)$ time to sort array of size n.

Theorem

Randomized Quick Sort sorts a given array of length n in $O(n \log n)$ expected time.

Note: On *every* input randomized quick sort takes $O(n \log n)$ time in expectation. On *every* input it may take $\Omega(n^2)$ time with some small probability.

Example: Verifying Matrix Multiplication

Problem

Given three $n \times n$ matrices A, B, C is $AB = C$?

Example: Verifying Matrix Multiplication

Problem

Given three $n \times n$ matrices A, B, C is $AB = C$?

Deterministic algorithm:

- Multiply A and B and check if equal to C.
- Running time?

Example: Verifying Matrix Multiplication

Problem

Given three $n \times n$ matrices A, B, C is $AB = C$?

Deterministic algorithm:

- Multiply A and B and check if equal to C.
- Running time? $O(n^3)$ by straight forward approach. $O(n^{2.37})$ with fast matrix multiplication (complicated and impractical).

Example: Verifying Matrix Multiplication

Problem

Given three $n \times n$ matrices A, B, C is $AB = C$?

Example: Verifying Matrix Multiplication

Problem

Given three $n \times n$ matrices A, B, C is $AB = C$?

Randomized algorithm:

- Pick a random $n \times 1$ vector r.
- Return the answer of the equality $ABr = Cr$.
- Running time?

Example: Verifying Matrix Multiplication

Problem

Given three $n \times n$ matrices A, B, C is $AB = C$?

Randomized algorithm:

- Pick a random $n \times 1$ vector r.
- Return the answer of the equality $ABr = Cr$.
- Running time? $O(n^2)$!

Example: Verifying Matrix Multiplication

Problem

Given three $n \times n$ matrices A, B, C is $AB = C$?

Randomized algorithm:

- Pick a random $n \times 1$ vector r.
- Return the answer of the equality $ABr = Cr$.
- Running time? $O(n^2)$!

Theorem

If $AB = C$ then the algorithm will always say YES. If $AB \neq C$ then the algorithm will say YES with probability at most $1/2$. Can repeat the algorithm 100 times to reduce the probability of false positive to $1/2^{100}$.

Why randomized algorithms?

- Many many applications in algorithms, data structures and computer science!
- In some cases only known algorithms are randomized or randomness is provably necessary.
- Often randomized algorithms are (much) simpler and/or more efficient.
- Several deep connections to mathematics, physics etc.
- ...
- Lots of fun!

Where do I get random bits?

Question: Are true random bits available in practice?

- Can use pseudo-random bits or semi-random bits from nature. Several fundamental unresolved questions in complexity theory on this topic. Beyong the scope of this course.
- In practice pseudo-random generators work quite well in many applications.
- The model is interesting to think in the abstract and is very useful even as a theoretical construct. One can *derandomize* randomized algorithms to obtain deterministic algorithms.

Average case analysis vs Randomized algorithms

Average case analysis:

- Fix a deterministic algorithm.
- Assume inputs comes from a probability distribution.
- Analyze the algorithm's *average* performance over the distribution over inputs.

Randomized algorithms:

- Algorithm uses random bits in addition to input.
- Analyze algorithms *average* performance over the given input where the average is over the random bits that the algorithm uses.
- On each input behaviour of algorithm is random. Analyze worst-case over all inputs of the (average) performance.

Discrete Probability

We restrict attention to finite probability spaces.

Definition

A discrete probability space is a pair (Ω, p) consists of finite set Ω of *elementary* events and function $p : \Omega \rightarrow [0, 1]$ which assigns a probability $p(\omega)$ for each $\omega \in \Omega$ such that $\sum_{\omega \in \Omega} p(\omega) = 1$.

Discrete Probability

We restrict attention to finite probability spaces.

Definition

A discrete probability space is a pair (Ω, p) consists of finite set Ω of *elementary* events and function $p : \Omega \rightarrow [0, 1]$ which assigns a probability $p(\omega)$ for each $\omega \in \Omega$ such that $\sum_{\omega \in \Omega} p(\omega) = 1$.

Example

An ubiased coin. $\Omega = \{H, T\}$ and $p(H) = p(T) = 1/2$.

Example

A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $p(i) = 1/6$ for $1 \leq i \leq 6$.

Example

A biased coin. $\Omega = \{H, T\}$ and $p(H) = 2/3, p(T) = 1/3$.

More Examples

Example

Two independent ubiased coins. $\Omega = \{HH, TT, HT, TH\}$ and $p(HH) = p(TT) = p(HT) = p(TH) = 1/4$.

Example

A pair of correlated dice. $\Omega = \{(i,j) \mid 1 \leq i \leq 6, 1 \leq j \leq 6\}$. $p(i,i) = 1/6$ for $1 \leq i \leq 6$ and $p(i,j) = 0$ if $i \neq j$.

Events

Definition

Given a probability space (Ω, p) an *event* is a subset of Ω. In other words an event is a collection of elementary events. The probability of an event A, denoted by $p(A)$, is $\sum_{\omega \in A} p(\omega)$. The complement of an event $A \subseteq \Omega$ is the event $\Omega \setminus A$ frequently denoted by $\bar{A}$.

Events

Definition

Given a probability space (Ω, p) an *event* is a subset of Ω. In other words an event is a collection of elementary events. The probability of an event A, denoted by $p(A)$, is $\sum_{\omega \in A} p(\omega)$. The complement of an event $A \subseteq \Omega$ is the event $\Omega \setminus A$ frequently denoted by $\bar{A}$.

Example

A pair of independent dice. $\Omega = \{(i,j) \mid 1 \leq i \leq 6, 1 \leq j \leq 6\}$.

- Let A be the event that the sum of the two numbers on the dice is even. Then $A = \{(i,j) \in \Omega \mid (i+j) \text{ is even}\}$.
  $p(A) = |A|/36 = 1/2$.
- Let B be the event that the first die has 1. Then
  $B = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)\}$.
  $p(B) = 6/36 = 1/6$.

Independent Events

Definition

Given a probability space (Ω, p) and two events A, B are *independent* if and only if $p(A \cap B) = p(A)p(B)$. Otherwise they are *dependent*. In other words A, B independent implies one does not affect the other.

Independent Events

Definition

Given a probability space (Ω, p) and two events A, B are *independent* if and only if $p(A \cap B) = p(A)p(B)$. Otherwise they are *dependent*. In other words A, B independent implies one does not affect the other.

Example

Two coins. $\Omega = \{HH, TT, HT, TH\}$ and $p(HH) = p(TT) = p(HT) = p(TH) = 1/4$.

- A is the event that the first coin is heads and B is the event that second coin is tails. A, B are independent.
- A is the event that the two coins are different. B is the event that the second coin is heads. A, B independent.
- A is the event that both are not tails and B is event that second coin is heads. A, B are dependent.

Random Variables

Definition

Given a probability space (Ω, p) a (real-valued) random variable X over Ω is a function that maps each elementary event to a real number. In other words $X : \Omega \rightarrow \mathbb{R}$.

Random Variables

Definition

Given a probability space (Ω, p) a (real-valued) random variable X over Ω is a function that maps each elementary event to a real number. In other words $X : \Omega \rightarrow \mathbb{R}$.

Example

A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $p(i) = 1/6$ for $1 \leq i \leq 6$.

- $X : \Omega \rightarrow \mathbb{R}$ where $X(i) = i \mod 2$.
- $Y : \Omega \rightarrow \mathbb{R}$ where $Y(i) = i^2$.

Random Variables

Definition

Given a probability space (Ω, p) a (real-valued) random variable X over Ω is a function that maps each elementary event to a real number. In other words $X : \Omega \rightarrow \mathbb{R}$.

Example

A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $p(i) = 1/6$ for $1 \leq i \leq 6$.

- $X : \Omega \rightarrow \mathbb{R}$ where $X(i) = i \mod 2$.
- $Y : \Omega \rightarrow \mathbb{R}$ where $Y(i) = i^2$.

Definition

A binary random variable is one that takes on values in $\{0, 1\}$.

Indicator Random Variables

Special type of random variables that are quite useful.

Definition

Given a probability space (Ω, p) and an event $A \subseteq \Omega$ the indicator random variable X_A is a binary random variable where $X_A(\omega) = 1$ if $\omega \in A$ and $X_A(\omega) = 0$ if $\omega \notin A$.

Indicator Random Variables

Special type of random variables that are quite useful.

Definition

Given a probability space (Ω, p) and an event $A \subseteq \Omega$ the indicator random variable X_A is a binary random variable where $X_A(\omega) = 1$ if $\omega \in A$ and $X_A(\omega) = 0$ if $\omega \notin A$.

Example

A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $p(i) = 1/6$ for $1 \leq i \leq 6$. Let A be the even that i is divisible by 3. Then $X_A(i) = 1$ if $i = 3, 6$ and 0 otherwise.

Expectation

Definition

For a random variable X over a probability space (Ω, p) the *expectation* of X is defined as $\sum_{\omega \in \Omega} p(\omega)X(\omega)$. In other words the average value of X according to the probabilities given by p.

Expectation

Definition

For a random variable X over a probability space (Ω, p) the *expectation* of X is defined as $\sum_{\omega \in \Omega} p(\omega)X(\omega)$. In other words the average value of X according to the probabilities given by p.

Example

A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $p(i) = 1/6$ for $1 \leq i \leq 6$.

- $X : \Omega \rightarrow \mathbb{R}$ where $X(i) = i \mod 2$. Then $E[X] = 1/2$.
- $Y : \Omega \rightarrow \mathbb{R}$ where $Y(i) = i^2$. Then $E[Y] = \sum_{i=1}^{6} \frac{1}{6} \cdot i^2 = 91/6$.

Expectation

Definition

For a random variable X over a probability space (Ω, p) the *expectation* of X is defined as $\sum_{\omega \in \Omega} p(\omega)X(\omega)$. In other words the average value of X according to the probabilities given by p.

Example

A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $p(i) = 1/6$ for $1 \leq i \leq 6$.

- $X : \Omega \rightarrow \mathbb{R}$ where $X(i) = i \mod 2$. Then $E[X] = 1/2$.
- $Y : \Omega \rightarrow \mathbb{R}$ where $Y(i) = i^2$. Then $E[Y] = \sum_{i=1}^{6} \frac{1}{6} \cdot i^2 = 91/6$

Proposition

For an indicator variable X_A, $E[X_A] = p(A)$.

Linearity of Expectation

Lemma

Let X, Y be two random variables over a probability space (Ω, p). Then $E[X + Y] = E[X] + E[Y]$.

Proof.

$$E[X + Y] = \sum_{\omega \in \Omega} p(\omega)(X(\omega) + Y(\omega))$$
$$= \sum_{\omega \in \Omega} p(\omega)X(\omega) + \sum_{\omega \in \Omega} p(\omega)Y(\omega) = E[X] + E[Y].$$

Linearity of Expectation

Lemma

Let X, Y be two random variables over a probability space (Ω, p). Then $E[X + Y] = E[X] + E[Y]$.

Proof.

$$E[X + Y] = \sum_{\omega \in \Omega} p(\omega)(X(\omega) + Y(\omega))$$
$$= \sum_{\omega \in \Omega} p(\omega)X(\omega) + \sum_{\omega \in \Omega} p(\omega)Y(\omega) = E[X] + E[Y].$$

□

Corollary

$E[a_1X_1 + a_2X_2 + \ldots + a_nX_n] = \sum_{i=1}^{n} a_iE[X_i]$.