Randomized Quick Sort

Fahad Panolan

Department of Computer Science and Engineering
Indian Institute of Technology Hyderabad, India

26-Jan-2021

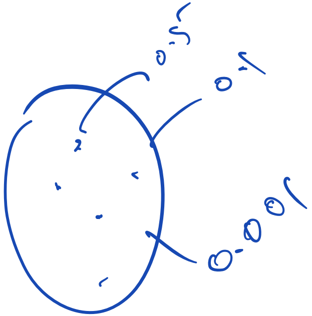
Basics of Discrete Probability

Discrete Probability

Definition

A discrete probability space is a pair $(\Omega, \Pr)$ where

- Ω is a countable set, called the set of elementary events.
- $\Pr : \Omega \to [0, 1]$ such that $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.



Discrete Probability

Definition

A discrete probability space is a pair $(\Omega, \Pr)$ where

- Ω is a countable set, called the set of elementary events.
- $\Pr : \Omega \to [0, 1]$ such that $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.

Examples

- An unbiased coin.

Discrete Probability

Definition

A discrete probability space is a pair $(\Omega, \Pr)$ where

- Ω is a countable set, called the set of elementary events.
- $\Pr : \Omega \to [0, 1]$ such that $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.

Examples

- An unbiased coin. $\Omega = \{H, T\}$ and $\Pr[H] = \Pr[T] = 1/2$.

Discrete Probability

Definition

A discrete probability space is a pair $(\Omega, \Pr)$ where

- Ω is a countable set, called the set of elementary events.
- $\Pr : \Omega \to [0, 1]$ such that $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.

Examples

- An unbiased coin. $\Omega = \{H, T\}$ and $\Pr[H] = \Pr[T] = 1/2$.
- A 6-sided unbiased die.

Discrete Probability

Definition

A discrete probability space is a pair $(\Omega, \Pr)$ where

- Ω is a countable set, called the set of elementary events.
- $\Pr : \Omega \to [0, 1]$ such that $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.

Examples

- An unbiased coin. $\Omega = \{H, T\}$ and $\Pr[H] = \Pr[T] = 1/2$.
- A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\Pr[i] = 1/6$ for all $i \in \Omega$.

Discrete Probability

Definition

A discrete probability space is a pair $(\Omega, \Pr)$ where

- Ω is a countable set, called the set of elementary events.
- $\Pr : \Omega \to [0,1]$ such that $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.

Examples

- An unbiased coin. $\Omega = \{H, T\}$ and $\Pr[H] = \Pr[T] = 1/2$.
- A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\Pr[i] = 1/6$ for all $i \in \Omega$.
- A pair of independent dice.

Discrete Probability

Definition

A discrete probability space is a pair $(\Omega, \Pr)$ where

- Ω is a countable set, called the set of elementary events.
- $\Pr : \Omega \to [0,1]$ such that $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.

Examples

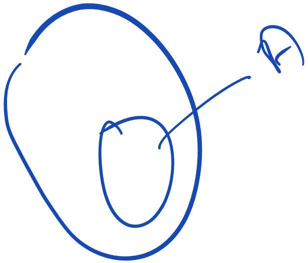
- An unbiased coin. $\Omega = \{H, T\}$ and $\Pr[H] = \Pr[T] = 1/2$.
- A 6-sided unbiased die. $\Omega = \{1,2,3,4,5,6\}$ and $\Pr[i] = 1/6$ for all $i \in \Omega$.
- A pair of independent dice. $\Omega = \{(i,j) \mid 1 \leq i \leq 6, 1 \leq j \leq 6\}$ and $\Pr[(i,j)] = 1/36$ for all $(i,j) \in \Omega$.

Events

Definition

Given a probability space $(\Omega, \Pr)$ an <u>event</u> is a subset of Ω. In other words an event is a collection of elementary events. The probability of an event A, denoted by $\Pr[A]$, is $\sum_{\omega \in A} \Pr[\omega]$.

The <u>complement event</u> of an event $A \subseteq \Omega$ is the event $\Omega \setminus A$ frequently denoted by $\bar{A}$.



Events

Definition

Given a probability space $(\Omega, \Pr)$ 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 $\Pr[A]$, is $\sum_{\omega \in A} \Pr[\omega]$.

The complement event 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 $A = \{(i,j) \in \Omega : (i+j) \text{ is even}\}$.
$\Pr[A] = |A|/36 = 1/2$.

Random Variables and Expectation

Random Variable

Given a probability space $(\Omega, \Pr)$ a random variable X over Ω is

$$X : \Omega \to \mathbb{R}.$$

Random Variables and Expectation

Random Variable

Given a probability space $(\Omega, \Pr)$ a random variable X over Ω is

$$X : \Omega \to \mathbb{R}.$$

Expectation

For a random variable X over a probability space $(\Omega, \Pr)$ the expectation of X is defined as

$$\sum_{\omega \in \Omega} \Pr[\omega] X(\omega).$$

In other words, the expectation is the average value of X according to the probabilities given by $\Pr[\cdot]$.

Expectation: examples

Example

A 6-sided unbiased die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\Pr[i] = 1/6$ for $1 \le i \le 6$.

- $X : \Omega \to \mathbb{R}$ where $X(i) = i \mod 2$. Then

$$\mathbf{E}[X] = \sum_{i=1}^{6} \Pr[i] \cdot X(i) = \frac{1}{6} \sum_{i=1}^{6} X(i) = 1/2.$$

$\frac{3}{6} = \frac{1}{2}$

Expectation: examples

Example

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

- $X : \Omega \to \mathbb{R}$ where $X(i) = i \mod 2$. Then

$$\mathbf{E}[X] = \sum_{i=1}^{6} \Pr[i] \cdot X(i) = \frac{1}{6} \sum_{i=1}^{6} X(i) = 1/2.$$

- $Y : \Omega \to \mathbb{R}$ where $Y(i) = i$. Then

$$\mathbf{E}[Y] = \sum_{i=1}^{6} \frac{1}{6} \cdot i = 3.5.$$

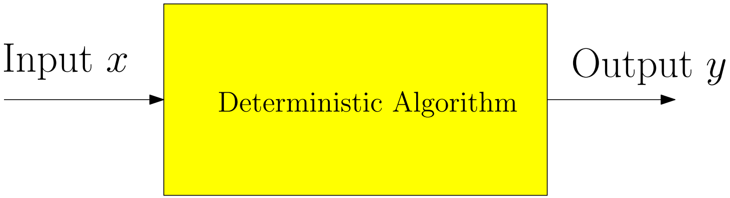
Suppose X and Y are two random variables over $(\Omega, \Pr)$.

Then $E[aX+bY] = aE[X]+bE[Y]$.

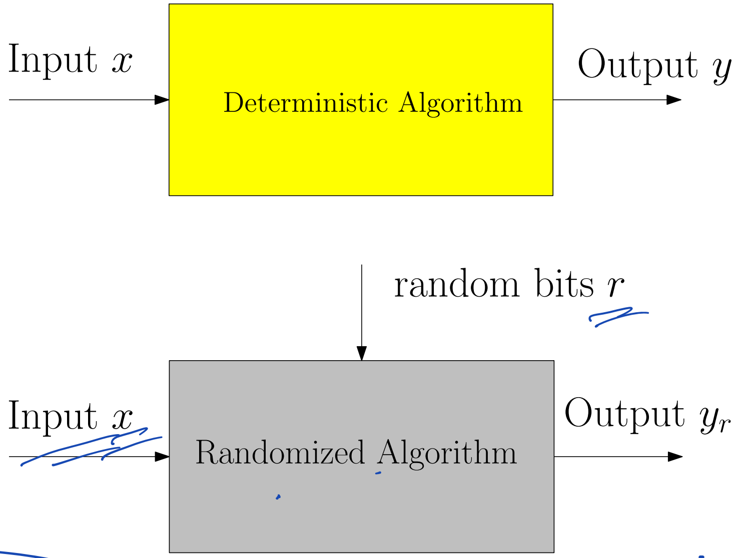
(Linearity of expectation).

Randomized Algorithms

Randomized Algorithms



Randomized Algorithms



Las Vegas , Monte Carlo.

Randomized Quick Sort

Ref: Slides of Chandra Chekuri

Run time $= O(\#\text{ comparisons})$

Quick Sort

- Pick a pivot element from array (last element as pivot)
- Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and equal to pivot.
- Recursively sort the subarrays, and concatenate them.

(Ω, P_r),

Randomized **Quick Sort**

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

The running time is random variable.

Analysis

Fix A

Let $Q(A)$ be number of comparisons done on input array A:

- For $1 \leq i < j < n$ let R_{ij} be the event that rank i element is compared with rank j element.
- X_{ij} is the indicator random variable for R_{ij}. That is, $X_{ij} = 1$ if rank i is compared with rank j element, otherwise 0.

$$E(Q(A)) = O(n \log n).$$

random choices

5	4	2	10	3	9

→

1	2	3	4	5	6
2	3	4	5	9	10

Analysis

Let $Q(A)$ be number of comparisons done on input array A:

- For $1 \le i < j < n$ let R_{ij} be the event that rank i element is compared with rank j element.
- X_{ij} is the indicator random variable for R_{ij}. That is, $X_{ij} = 1$ if rank i is compared with rank j element, otherwise 0.

$$Q(A) = \sum_{1 \le i < j \le n} X_{ij}$$

and hence by linearity of expectation,

$$E[Q(A)] = \sum_{1 \le i < j \le n} E[X_{ij}] = \sum_{1 \le i < j \le n} \Pr[R_{ij}].$$

Analysis

$\boxed{R_{ij} = \text{rank } i \text{ element is compared with rank } j \text{ element.}}$

Question: What is $\Pr[R_{ij}]$?

Analysis

$\boxed{R_{ij} = \text{rank } i \text{ element is compared with rank } j \text{ element.}}$

Question: What is $\Pr[R_{ij}]$?

	7	5	9	1	3	4	8	6
Ranks:	6	4	8	1	2	3	7	5

Analysis

$\boxed{R_{ij} = \text{rank } i \text{ element is compared with rank } j \text{ element.}}$

Question: What is $\Pr[R_{ij}]$?

7	5	9	1	3	4	8	6
6	4	8	1	2	3	7	5

Ranks: 6 4 8 1 2 3 7 5

Probability of comparing 5 to 8 is $\Pr[R_{4,7}]$.

Analysis

$\boxed{R_{ij} = \text{rank } i \text{ element is compared with rank } j \text{ element.}}$

Question: What is $\Pr[R_{ij}]$?

7	5	9	1	3	4	8	6
6	4	8	1	2	3	7	5

Ranks: 6 4 8 1 2 3 7 5

Probability of comparing 5 to 8 is $\Pr[R_{4,7}]$.

- If pivot too small (say 3 [rank 2]). Partition and call recursively:

7	5	9	1	3	4	8	6

$\Longrightarrow$

1		3		7	5	9	4	8	6

Decision if to compare 5 to 8 is moved to subproblem.

- If pivot too large (say 9 [rank 8]):

7	5	9	1	3	4	8	6

$\Longrightarrow$

7	5	1	3	4	8	6		9

Decision if to compare 5 to 8 moved to subproblem.

Analysis

- If pivot is 5 (rank 4).

7	5	9	1	3	4	8	6

$\Longrightarrow$

1	3	4		5		7	9	8	6

- If pivot is 8 (rank 7).

7	5	9	1	3	4	8	6

$\Longrightarrow$

7	5	1	3	4	6		8		9

- If pivot in between the two numbers (say 6 [rank 5]):

7	5	9	1	3	4	8	6

$\Longrightarrow$

5	1	3	4		6		7	8	9

5 and 8 will never be compared to each other.

Analysis

Question: What is $\Pr[R_{i,j}]$?

Conclusion

$R_{i,j}$ happens if and only if :

$$\boxed{i\text{th or } j\text{th ranked element is the first pivot out of the elements ranked } i \text{ to } j.}$$

4 5 6 7

Analysis

Question: What is $\Pr[R_{ij}]$?

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

Analysis

Question: What is $\Pr[R_{ij}]$?

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

Analysis

Question: What is $\Pr[R_{ij}]$?

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

Proof.

Let $a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n$ be elements of A in sorted order. Let $S = \{a_i, a_{i+1}, \ldots, a_j\}$

Observation: If pivot is chosen outside S then all of S either in left array or right array.

Observation: a_i and a_j separated when a pivot is chosen from S for the first time. Once separated no comparison.

Observation: a_i is compared with a_j if and only if either a_i or a_j is chosen as a pivot from S at separation. □

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

Proof.

Let $a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n$ be sort of A. Let $S = \{a_i, a_{i+1}, \ldots, a_j\}$
Observation: a_i is compared with a_j if and only if either a_i or a_j is chosen as a pivot from S at separation.
Observation: Given that pivot is chosen from S the probability that it is a_i or a_j is exactly $\frac{2}{|S|} = \frac{2}{(j-i+1)}$ since the pivot is chosen uniformly at random from the array. $\square$

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

$$\mathbf{E}[Q(A)] = \sum_{1 \le i < j \le n} \Pr[R_{ij}]$$

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

$$\mathbf{E}[Q(A)] = \sum_{1 \le i < j \le n} \Pr[R_{ij}] = \sum_{1 \le i < j \le n} \frac{2}{j-i+1}$$

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

$$\begin{aligned}
\mathbf{E}[Q(A)] = \sum_{1 \le i < j \le n} \Pr[R_{ij}] &= \sum_{1 \le i < j \le n} \frac{2}{j-i+1} \\
&= 2 \sum_{i=1}^{n-1} \sum_{i<j}^{n} \frac{1}{j-i+1}
\end{aligned}$$

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

$$
\begin{aligned}
\mathbf{E}[Q(A)] = \sum_{1 \le i < j \le n} \Pr[R_{ij}] \quad &= \quad \sum_{1 \le i < j \le n} \frac{2}{j-i+1} \\
&= \quad 2 \sum_{i=1}^{n-1} \sum_{i<j}^{n} \frac{1}{j-i+1} \\
&\le \quad 2 \sum_{i=1}^{n-1} \sum_{\Delta=2}^{n-i+1} \frac{1}{\Delta}
\end{aligned}
$$

Analysis

Lemma

$\Pr[R_{ij}] = \frac{2}{j-i+1}$.

$$\begin{aligned}
\mathbf{E}[Q(A)] = \sum_{1 \le i < j \le n} \Pr[R_{ij}] &= \sum_{1 \le i < j \le n} \frac{2}{j-i+1} \\
&= 2 \sum_{i=1}^{n-1} \sum_{i<j}^{n} \frac{1}{j-i+1} \\
&\le 2 \sum_{i=1}^{n-1} \sum_{\Delta=2}^{n-i+1} \frac{1}{\Delta} \\
&\le 2 \sum_{1 \le i < n} H_n
\end{aligned}$$

$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots \frac{1}{n}.$

Analysis

$$\le 2 \sum c_2 \log n$$
$$\le 2c_2 n \log n = \Theta(n \log n)$$

$H_n = \sum_{i=1}^{n} \frac{1}{i}$ is the n'th harmonic number

1. $H_n = \Theta(1)$.
2. $H_n = \Theta(\log \log n)$.
3. $H_n = \Theta(\sqrt{\log n})$.
4. $H_n = \Theta(\log n)$. ✓
5. $H_n = \Theta(\log^2 n)$.

Exc

$\exists\ c_1$ and c_2

s.t. $c_1 \log n \le H_n \le c_2 \log n$

Analysis

Analysis

Theorem

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

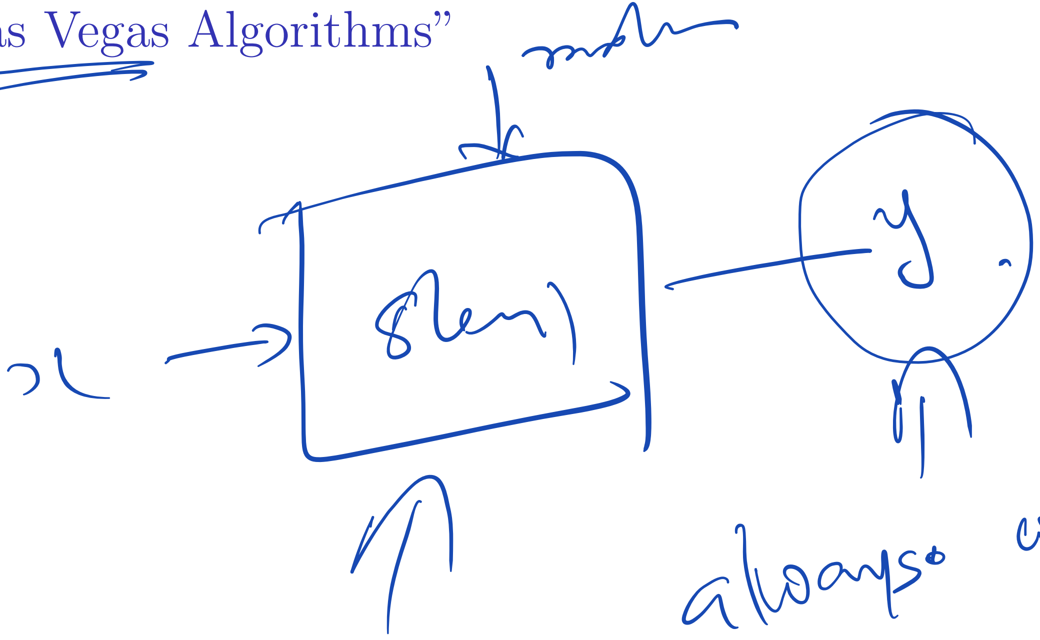
Analysis

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.

Analyzing "Las Vegas Algorithms"



always correct.

running is a random variable.

Analyzing "Las Vegas Algorithms"

Randomized algorithm $\mathcal{A}$ for a problem Π:

- Let $R(x)$ be the time for $\mathcal{A}$ to run on input x of length $|x|$.

Analyzing "Las Vegas Algorithms"

Randomized algorithm $\mathcal{A}$ for a problem Π:

- Let $R(x)$ be the time for $\mathcal{A}$ to run on input x of length $|x|$.
- $R(x)$ is a random variable: depends on random bits used by $\cancel{R}$ $\mathcal{A}$.

Analyzing “Las Vegas Algorithms”

Randomized algorithm $\mathcal{A}$ for a problem Π:

- Let $R(x)$ be the time for $\mathcal{A}$ to run on input x of length $|x|$.
- $R(x)$ is a random variable: depends on random bits used by R.
- $\mathbf{E}[R(x)]$ is the expected running time for ~~R~~ A on x.

Analyzing "Las Vegas Algorithms"

Randomized algorithm $\mathcal{A}$ for a problem Π:

- Let $R(x)$ be the time for $\mathcal{A}$ to run on input x of length $|x|$.
- $R(x)$ is a random variable: depends on random bits used by R.
- $\mathbf{E}[R(x)]$ is the expected running time for R on x
- Expected time on worst input of size n

$$T_{rand}(n) = \max_{x:|x|=n} \mathbf{E}[R(x)].$$

$O(\log n)$,

For R. QuickSort:

$O(n \log n)$

Thank You.