CS 573: Algorithms, Fall 2014

Randomized Algorithms

Lecture 9

September 23, 2014

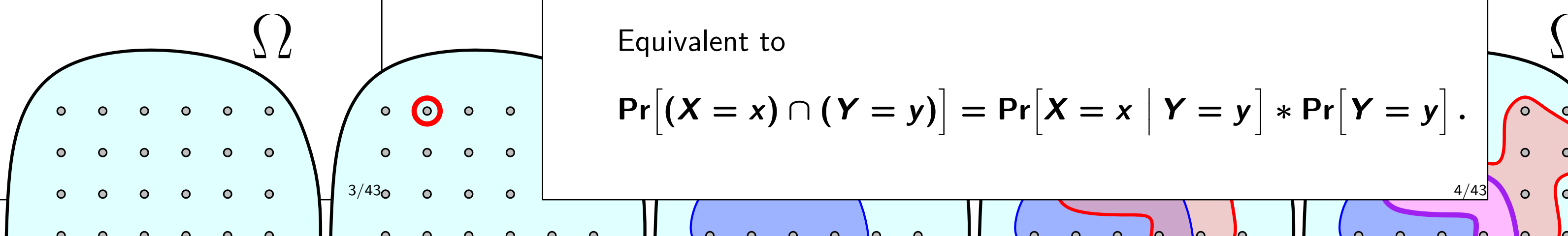
Part I

Randomized Algorithms

Probability - quick review

With pictures

1. Ω: Sample space
2. Ω: Is a set of ***elementary event***/***atomic event***/***simple event***.
3. Every atomic event $x \in \Omega$ has ***Probability*** $\Pr[x]$.
4. $X \equiv f(x)$: Random variable associate a value with each atomic event $x \in \Omega$.
5. $\mathbf{E}[X]$: ***Expectation***: The average value of the random variable $X \equiv f(x)$. $\mathbf{E}[X] = \sum_{x \in X} f(x) * \Pr[X = x]$.
6. An event $A \subseteq \Omega$ is a collection of atomic events.



Probability - quick review

Definitions

Definition (Informal)

Random variable: a function from probability space to $\mathbb{R}$. Associates value $\forall$ atomic events in probability space.

Definition

The ***conditional probability*** of X given Y is

$$\Pr\left[X = x \,\middle|\, Y = y\right] = \frac{\Pr\left[(X = x) \cap (Y = y)\right]}{\Pr\left[Y = y\right]}.$$

Equivalent to

$$\Pr\left[(X = x) \cap (Y = y)\right] = \Pr\left[X = x \,\middle|\, Y = y\right] * \Pr\left[Y = y\right].$$

Probability - quick review

Even more definitions

Definition

The events $X = x$ and $Y = y$ are ***independent***, if

$$\Pr[X = x \cap Y = y] = \Pr[X = x] \cdot \Pr[Y = y] .$$
$$\equiv \Pr\left[X = x \mid Y = y\right] = \Pr[X = x] .$$

Definition

The ***expectation*** of a random variable X its average value:

$$\mathbf{E}\left[X\right] = \sum_x x \cdot \Pr[X = x] ,$$

Linearity of expectations

Lemma (Linearity of expectation.)

$\forall$ random variables X and Y: $\mathbf{E}\left[X + Y\right] = \mathbf{E}\left[X\right] + \mathbf{E}\left[Y\right]$.

Proof.

Use definitions, do the math. See notes for details. □

Probability - quick review

Conditional Expectation

Definition

X, Y: random variables. The ***conditional expectation*** of X given Y (i.e., you know $Y = y$):

$$\mathbf{E}\left[X \mid Y\right] = \mathbf{E}\left[X \mid Y = y\right] = \sum_x x * \Pr\left[X = x \mid Y = y\right] .$$

$\mathbf{E}[X]$ is a number.
$f(y) = \mathbf{E}\left[X \mid Y = y\right]$ is a function.

Conditional Expectation

Lemma

$\forall$ X, Y (not necessarily independent): $\mathbf{E}[X] = \mathbf{E}\left[\mathbf{E}\left[X \mid Y\right]\right]$.

$$\mathbf{E}\left[\mathbf{E}\left[X \mid Y\right]\right] = \mathbf{E}_y\left[\mathbf{E}\left[X \mid Y = y\right]\right]$$

Proof.

Use definitions, and do the math. See class notes. □

Sorting Nuts & Bolts

Problem (**Sorting Nuts and Bolts**)

1. *Input: Set n nuts + n*

Sorting nuts & bolts...

Algorithm

1. Naive algorithm...
2. ...better algorithm?

Sorting nuts & bolts...

```
MatchNutsAndBolts(N: nuts, B: bolts)
    Pick a random nut n_pivot from N
    Find its matching bolt b_pivot in B
    B_L ← All bolts in B smaller than n_pivot
    N_L ← All nuts in N smaller than b_pivot
    B_R ← All bolts in B larger than n_pivot
    N_R ← All nuts in N larger than b_pivot
    MatchNutsAndBolts(N_R, B_R)
    MatchNutsAndBolts(N_L, B_L)
```

QuickSort style...

What is running time for randomized algorithms?

Definitions

Definition

$\mathcal{RT}(U)$: random variable – ***running time*** of the algorithm on input U.

Definition

Expected running time $\mathbf{E}[\mathcal{RT}(U)]$ for input U.

Definition

expected running-time of algorithm for input size n:

$$T(n) = \max_{U \text{ is an input of size } n} \mathbf{E}\Big[\mathcal{RT}(U)\Big].$$

What is running time for randomized algorithms?

More definitions

Definition

$\mathrm{rank}(x)$: ***rank*** of element $x \in S$ = number of elements in S smaller or equal to x.

Nuts and bolts running time

Theorem

Expected running time **MatchNutsAndBolts** (**QuickSort**) *is* $T(n) = O(n \log n)$. *Worst case is* $O(n^2)$.

Proof.

$\Pr[\mathrm{rank}(n_{pivot}) = k] = \frac{1}{n}$. Thus,

$$\begin{aligned} T(n) &= \mathop{\mathbf{E}}_{k=\mathrm{rank}(n_{pivot})}\Big[O(n) + T(k-1) + T(n-k)\Big] \\ &= O(n) + \mathop{\mathbf{E}}_{k}[T(k-1) + T(n-k)] \\ &= O(n) + \sum_{k=1}^{n} \Pr[Rank(Pivot) = k] \\ &\qquad\qquad * (T(k-1) + T(n-k)) \\ &= O(n) + \sum_{k=1}^{n} \frac{1}{n} \cdot (T(k-1) + T(n-k)), \end{aligned}$$

Solution is $T(n) = O(n \log n)$. □

Alternative intuitive analysis...

Which is not formally correct

1. **MatchNutsAndBolts** is ***lucky*** if $\frac{n}{4} \leq \mathrm{rank}(n_{pivot}) \leq \frac{3}{4}n$.
2. $\Pr[\text{"lucky"}] = 1/2$.
3. $T(n) \leq O(n) + \Pr[\text{"lucky"}] * (T(n/4) + T(3n/4)) + \Pr[\text{"unlucky"}] * T(n)$.
4. $T(n) = O(n) + \frac{1}{2} * \left(T(\frac{n}{4}) + T(\frac{3}{4}n)\right) + \frac{1}{2} T(n)$.
5. Rewriting: $T(n) = O(n) + T(n/4) + T((3/4)n)$.
6. ... solution is $O(n \log n)$.

Worst case vs. average case

Expected running time of a randomized algorithm is

$$T(n) = \max_{U \text{ is an input of size } n} \mathbf{E}\Big[\mathcal{RT}(U)\Big],$$

Worst case running time of deterministic algorithm:

$$T(n) = \max_{U \text{ is an input of size } n} \mathcal{RT}(U),$$

High Probability running time...

Definition

Running time **Alg** is $O(f(n))$ with ***high probability*** if

$$\Pr\left[\mathcal{RT}(\mathsf{Alg}(n)) \geq c \cdot f(n)\right] = o(1).$$

$$\implies \Pr\left[\mathcal{RT}(\mathsf{Alg}) > c * f(n)\right] \to 0 \text{ as } n \to \infty.$$

Usually use weaker def:

$$\Pr\left[\mathcal{RT}(\mathsf{Alg}(n)) \geq c \cdot f(n)\right] \leq \frac{1}{n^d},$$

Technical reasons... also assume that
$\mathbf{E}[\mathcal{RT}(\mathsf{Alg}(n))] = O(f(n))$.

Part II

Slick analysis of QuickSort

A Slick Analysis of **QuickSort**

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

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

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

and hence by linearity of expectation,

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

A Slick Analysis of **QuickSort**

R_{ij} = rank i element is compared with rank j element.

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

7	5	9	1	3	4	8	6

With ranks:

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

As such, probability of comparing **5** to **8** is $\Pr[R_{4,7}]$.

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

   7 5 9 1 (3) 4 8 6 $\implies$

   1 | 3 | 7 5 9 4 8 6

   Decision if to compare **5** to **8** is moved to subproblem.
2. If pivot too large (say **9** [rank 8]):

   7 5 (9) 1 3 4 8 6 $\implies$ 7 5 1 3 4 8 6 | 9

A Slick Analysis of QuickSort

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

6 4 8 1 2 3 7 5

As such, probability of comparing **5** to **8** is $\mathbf{Pr}[R_{4,7}]$.

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

1. If pivot is **5** (rank 4). Bingo!

   7 | 5 | 9 | 1 | 3 | 4 | 8 | 6 $\Longrightarrow$

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

2. If pivot is **8** (rank 7). Bingo!

   7 | 5 | 9 | 1 | 3 | 4 | 8 | 6 $\Longrightarrow$

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

3. 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.

A Slick Analysis of QuickSort

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

Conclusion:

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

> ith or jth ranked element is the first pivot out of ith to jth ranked elements.

How to analyze this?

Thinking acrobatics!

1. Assign every element in the array a random priority (say in $[0, 1]$).
2. Choose pivot to be the element with lowest priority in subproblem.
3. Equivalent to picking pivot uniformly at random (as **QuickSort** do).

A Slick Analysis of QuickSort

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

How to analyze this?

Thinking acrobatics!

1. Assign every element in the array a random priority (say in $[0, 1]$).
2. Choose pivot to be the element with lowest priority in subproblem.

$\Longrightarrow$ $R_{i,j}$ happens if either i or j have lowest priority out of elements rank i to j,

There are $k = j - i + 1$ relevant elements.

$$\mathbf{Pr}\Big[R_{i,j}\Big] = \frac{2}{k} = \frac{2}{j - i + 1}.$$

A Slick Analysis of QuickSort

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

Lemma

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

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... □

A Slick Analysis of **QuickSort**

Continued...

Lemma

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

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 $2/|S| = 2/(j-i+1)$ since the pivot is chosen uniformly at random from the array. □

A Slick Analysis of **QuickSort**

Continued...

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

Lemma

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

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

Part III

Quick Select

Randomized Quick Selection

Input Unsorted array A of n integers

Goal Find the jth smallest number in A (*rank j* number)

Randomized Quick Selection

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. Return pivot if rank of pivot is j.
4. Otherwise recurse on one of the arrays depending on j and their sizes.

Algorithm for Randomized Selection

Assume for simplicity that A has distinct elements.

```
QuickSelect(A, j):
    Pick pivot x uniformly at random from A
    Partition A into A_less, x, and A_greater using x as pivot
    if (|A_less| = j - 1) then
        return x
    if (|A_less| >= j) then
        return QuickSelect(A_less, j)
    else
        return QuickSelect(A_greater, j - |A_less| - 1)
```


QuickSelect analysis

1. $S_1, S_2, \ldots, S_k$ be the subproblems considered by the algorithm. Here $|S_1| = n$.
2. S_i would be ***successful*** if $|S_i| \leq (3/4)\,|S_{i-1}|$
3. Y_1 = number of recursive calls till first successful iteration. Clearly, total work till this happens is $O(Y_1 n)$.
4. n_i = size of the subproblem immediately after the $(i-1)$th successful iteration.
5. Y_i = number of recursive calls after the $(i-1)$th successful call, till the ith successful iteration.
6. Running time is $O(\sum_i n_i Y_i)$.

QuickSelect analysis

Example

S_i = subarray used in ith recursive call

$|S_i|$ = size of this subarray

Red indicates successful iteration.

Inst'	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
$\lvert S_i \rvert$	100	70	60	50	40	30	25	5	2
Succ'	$Y_1 = 2$		$Y_2 = 4$				$Y_3 = 2$		$Y_4 = 1$
$n_i =$	$n_1 = 100$		$n_2 = 60$				$n_3 = 25$		$n_4 = 2$

1. All the subproblems after $(i-1)$th successful iteration till ith successful iteration have size $\leq n_i$.
2. Total work: $O(\sum_i n_i Y_i)$.

QuickSelect analysis

Total work: $O(\sum_i n_i Y_i)$.

We have:

1. $n_i \leq (3/4) n_{i-1} \leq (3/4)^{i-1} n$.
2. Y_i is a random variable with geometric distribution Probability of $Y_i = k$ is $1/2^i$.
3. $\mathbf{E}[Y_i] = 2$.

As such, expected work is proportional to

$$\mathbf{E}\left[\sum_i n_i Y_i\right] = \sum_i \mathbf{E}\left[n_i Y_i\right] \leq \sum_i \mathbf{E}\left[(3/4)^{i-1} n Y_i\right]$$
$$= n \sum_i (3/4)^{i-1}\,\mathbf{E}\left[Y_i\right] = n \sum_{i=1} (3/4)^{i-1} 2 \leq 8n.$$

QuickSelect analysis

Theorem

The expected running time of **QuickSelect** *is* $O(n)$.