Chapter 6
Randomization

Algorithm Theory
WS 2014/15

Contention Resolution
Primality Test

Fabian Kuhn

Randomization

Randomized Algorithm:

- An algorithm that uses (or can use) random coin flips in order to make decisions

We will see: randomization can be a powerful tool to

- Make algorithms faster
- Make algorithms simpler
- Make the analysis simpler
  - Sometimes it's also the opposite…
- Allow to solve problems (efficiently) that cannot be solved (efficiently) without randomization
  - True in some computational models (e.g., for distributed algorithms)
  - Not clear in the standard sequential model

Randomized Quicksort

→ choose pivot at random

Quicksort:

S	v

$S_\ell < v$	v	$S_r > v$

function Quick (S: sequence): sequence;

{returns the sorted sequence S}

begin

  if $\#S \leq 1$ then **return** S

  else { choose pivot element v in S;

    partition S into S_ℓ with elements $< v$,

    and S_r with elements $> v$

    return $\boxed{\text{Quick}(S_\ell)}\boxed{v}\boxed{\text{Quick}(S_r)}$

end;

Randomized Quicksort Analysis

Randomized Quicksort: pick uniform random element as pivot

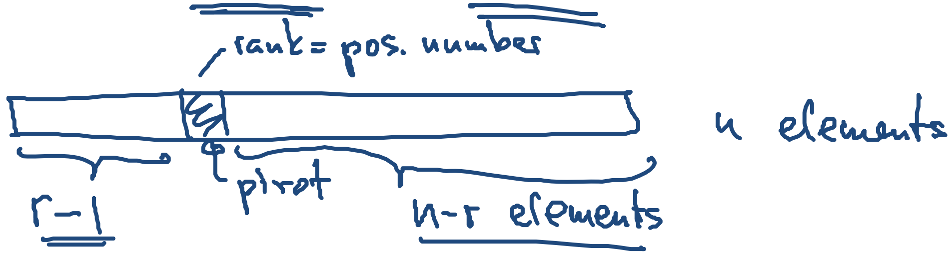
Running Time of sorting n elements:

- Let's just count the number of comparisons
- In the partitioning step, all $n-1$ non-pivot elements have to be compared to the pivot

- **Number of comparisons:**

$$\underline{\boldsymbol{n-1}} + \underline{\textbf{\#comparisons in recursive calls}}$$

- **If rank of pivot is $\boldsymbol{r}$:**
recursive calls with $\boldsymbol{r-1}$ and $\boldsymbol{n-r}$ elements



Randomized Quicksort Analysis

Random variables:

- C: total number of comparisons (for a given array of length n)
- R: rank of first pivot
- C_ℓ, C_r: number of comparisons for the 2 recursive calls

$$C = n - 1 + C_\ell + C_r$$

Expectation:

$$\mathbb{E}[C] = \mathbb{E}[n - 1 + C_\ell + C_r]$$

Linearity of Expectation:

$$\mathbb{E}[C] = n - 1 + \mathbb{E}[C_\ell] + \mathbb{E}[C_r]$$

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Randomized Quicksort Analysis

Random variables:

- C: total number of comparisons (for a given array of length n)
- R: rank of first pivot
- C_ℓ, C_r: number of comparisons for the 2 recursive calls

$$\boxed{\mathbb{E}[C] = n - 1 + \mathbb{E}[C_\ell] + \mathbb{E}[C_r]}$$

Law of Total Expectation:

$$E[X] = \sum_y P(Y=y) \cdot E[X \mid Y=y]$$

$$P(X=x) = \sum_y P(Y=y) \cdot P(X=x \mid Y=y)$$

$$\mathbb{E}[C] = \sum_{r=1}^{n} \mathbb{P}(R = r) \cdot \mathbb{E}[C \mid R = r]$$

n elements

$$= \sum_{r=1}^{n} \mathbb{P}(R = r) \cdot \big(n - 1 + \underbrace{\mathbb{E}[C_\ell \mid R = r]}_{r-1 \text{ elements}} + \underbrace{\mathbb{E}[C_r \mid R = r]}_{n-r \text{ elements}}\big)$$

Randomized Quicksort Analysis

We have seen that:

$$\mathbb{E}[C] = \sum_{r=1}^{n} \underbrace{\mathbb{P}(R = r)}_{\frac{1}{n}} \cdot (n - 1 + \mathbb{E}[C_\ell | R = r] + \mathbb{E}[C_r | R = r])$$

Define:

- $\boldsymbol{T(n)}$: expected number of comparisons when sorting n elements

$$\mathbb{E}[C] = T(n)$$
$$\mathbb{E}[C_\ell | R = r] = T(r-1)$$
$$\mathbb{E}[C_r | R = r] = T(n-r)$$

Recursion:

$\mathbb{P}(R=r)$

$$\boldsymbol{T(n) = \sum_{r=1}^{n} \frac{1}{n} \cdot \big(n - 1 + T(r-1) + T(n-r)\big)}$$

$$\boldsymbol{T(0) = T(1) = 0}$$

Randomized Quicksort Analysis

Theorem: The expected number of comparisons when sorting n elements using randomized quicksort is $\underline{T(n)} \le \underline{2n\ln n}$.

Proof:

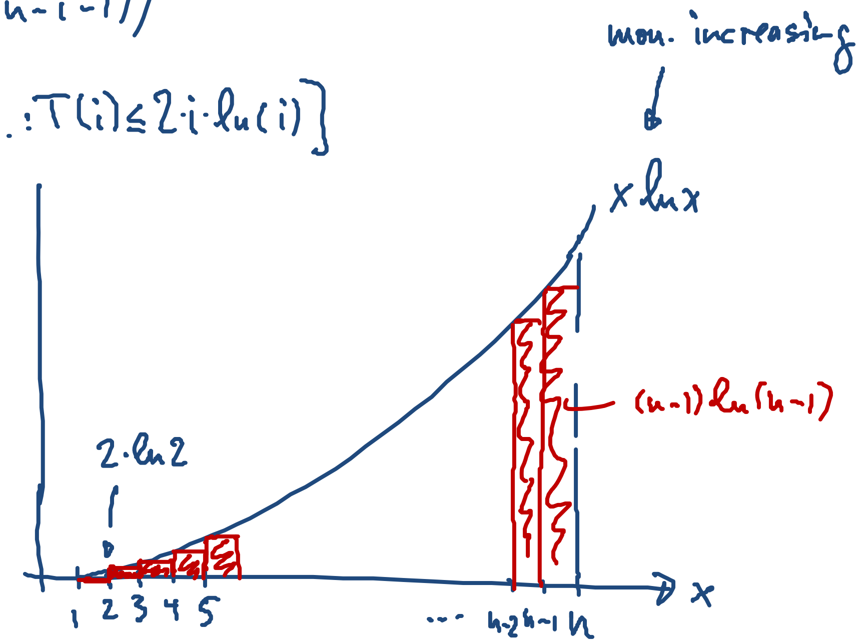
$$T(n) = \sum_{r=1}^{n} \frac{1}{n}\cdot\left(\underline{n-1} + T(\overbrace{r-1}^{=i}) + T(\overbrace{n-r}^{=n-i-1})\right), \qquad \underline{T(1)=0}$$

$$= n-1 + \frac{1}{n}\cdot\sum_{i=0}^{n-1}\left(T(i) + T(n-i-1)\right)$$

$$= n-1 + \frac{2}{n}\sum_{i=1}^{n-1} T(i) \qquad [\text{I.H.}: T(i) \le 2i\cdot\ln(i)]$$

$$\le n-1 + \frac{4}{n}\cdot\underline{\sum_{i=1}^{n-1} i\cdot\ln(i)}$$

$$\le n-1 + \frac{4}{n}\cdot\int_1^n x\ln(x)\,dx$$



Randomized Quicksort Analysis

Theorem: The expected number of comparisons when sorting n elements using randomized quicksort is $T(n) \leq 2n \ln n$.

Proof:

$$T(n) \leq n - 1 + \frac{4}{n} \cdot \int_1^n x \ln x \, dx$$

$$\int x \ln x \, dx = \frac{x^2 \ln x}{2} - \frac{x^2}{4}$$

$$T(n) \leq n-1 + \frac{4}{n}\left(\frac{n^2 \ln n}{2} - \frac{n^2}{4} + \frac{1}{4}\right)$$

$$= n - 1 + 2n\ln n - n + \frac{1}{n}$$

$$= 2n \ln n + \underbrace{\frac{1}{n} - 1}_{\leq 0} \leq 2n\ln(n)$$

$$\hookrightarrow E[C] \leq 2 \cdot n \cdot \ln(n)$$

also possible to show that

$T(n) = O(n \log n)$ with high prob.

(with prob. $1 - \frac{1}{n^c}$)