CSE 102 5-2-24

Quicksort (7.1 - 7.4)

sort $A[p..r]$ in increasing order

Quicksort(A)

1. if $p < r$
2. $q = \text{Partition}(A, p, r)$
3. $\text{Quicksort}(A, p, q-1)$
4. $\text{Quicksort}(A, q+1, r)$

note

Partition(A, p, r) re-aranges $A[p..r]$ so that

$$A[p\cdots(q-1)] \leq \underbrace{A[q]}_{\uparrow} < A[(q+1)\cdots r]$$

Pivot element

Pivot index: q

then returns q.

How does Partition work?

First! Pick Pivot to be $A[r]$

Partition (A, p, r)

```
i = p-1
for j = p to (r-1)
        if A[j] ≤ A[r]
                i += 1
                A[i] <-> A[j]   // swap
A[i+1] <-> A[r]   // swap
return (i+1)
```

Invariants:

$$\underbrace{A_p \cdots A_i}_{\leq \text{Pivot}} \quad \underbrace{A_{i+1} \cdots A_{j-1}}_{> \text{Pivot}} \quad \underbrace{A_j \cdots A_{r-1}}_{\text{unknown}} \quad \underbrace{A_r}_{\text{Pivot}}$$

if $A_j \le A_r$: swap $A_j \leftrightarrow A_{i+1}$, i++, j++

if $A_j > A_r$: j++

Exercise

Prove correctness of Quicksort() by induction. (Assume correctness of Partition()).

Runtime

Pivot: Same in best, worst, avg case.

$\#comp = (r-p+1)-1 = r-p$

at the level : $\#comp = n-1$

Quicksort!

- Worst case: $\Theta(n^2)$
- Avg. case: $\Theta(n \log n)$

Worst case!

occurs when $A[1 \dots n]$ is already sorted. Partition$(A, 1, n)$ gives

$$A[1 \dots (n-1)] \le \underbrace{A[n]}_{\substack{\text{pivot} \\ q=n}} < \dots \text{empty} \dots$$

so

$$T(n) = \begin{cases} 0 & \text{if } n = 0, 1 \\ T(n-1) + (n-1) & \text{if } n \ge 2 \end{cases}$$

By iteration:

$$T(n) = (n-1) + T(n-1)$$

$$= (n-1) + (n-2) + T(n-2)$$

$$= (n-1) + (n-2) + (n-3) + T(n-3)$$

$$\vdots$$

$$= \sum_{i=1}^{k} (n-i) + T(n-k)$$

halt when $n-k=1$, i.e. $\boxed{k = n-1}$

$$\therefore T(n) = \sum_{i=1}^{n-1} (n-i) + 0$$

$$= \sum_{j=1}^{n-1} n - \sum_{i=1}^{n-1} i$$

$$= n(n-1) - \frac{n(n-1)}{2} \quad \#$$

so $\boxed{T(n) = \frac{1}{2}n(n-1)}$ ← exact soln.

asym. soln: $\boxed{T(n) = \Theta(n^2)}$

Average case:

Assume all $n!$ permutations of $A[1..n]$ are equally likely. Let

$t(n)$ = average # of comparisons by Quicksort on arrays of len. n.

$$t(n) = \frac{\displaystyle\sum_{\text{all permutations}} \left(\begin{array}{c}\#\text{ comp. by Q.S.}\\ \text{on given perm.}\end{array}\right)}{n!}$$

<u>note</u>

Our assumption implies that the pivot is equally likely to be any element in $A[1..n]$, so q is eq. likely to be any # $1, 2, \ldots, n$.

<u>also</u>

- Partition$(A,1,n)$ does $n-1$ comparisons.

<u>also</u>

- $\text{len}\left(A[1\ldots(q-1)]\right) = q-1$
- $\text{len}\left(A[(q+1)\ldots n]\right) = n-q$

Then

$$t(n) = \sum_{q=1}^{n} \left((n-1) + t(q-1) + t(n-q)\right) \cdot \frac{1}{n}$$

$$= \sum_{q=1}^{n} (n-1) \cdot \frac{1}{n} + \sum_{q=1}^{n} t(q-1) \cdot \frac{1}{n} + \sum_{q=1}^{n} t(n-q) \cdot \frac{1}{n}$$

(use $t(0)=0$)

$$= (n-1) + \frac{1}{n}\left(\underbrace{\sum_{q=2}^{n} t(q-1)}_{q \leftarrow q+1} + \sum_{q=1}^{n-1} t(n-q)\right)$$

$$= (n-1) + \frac{1}{n}\left(\sum_{q=1}^{n-1} t(q) + \sum_{q=1}^{n-1} t(n-q)\right)$$

$$\therefore \boxed{t(n) = (n-1) + \frac{2}{n} \cdot \sum_{q=1}^{n-1} t(q)}$$

Let $X_n = \sum_{q=1}^{n-1} t(q)$, $X_1 = 0$

Then

$$X_{n+1} - X_n = \sum_{q=1}^{n} t(q) - \sum_{q=1}^{n-1} t(q) = t(n)$$

so

$$X_{n+1} - X_n = (n-1) + \frac{2}{n} \cdot X_n$$

$$\therefore \quad X_{n+1} - \left(\frac{n+2}{n}\right) X_n = n-1$$

multiply by magic #! $\dfrac{1}{(n+1)(n+2)}$

$$\therefore \quad \frac{X_{n+1}}{(n+1)(n+2)} - \frac{X_n}{n(n+1)} = \frac{n-1}{(n+1)(n+2)}$$

$$'' = \frac{3}{n+2} - \frac{2}{n+1}$$

Replace n by k:

$$\frac{x_{k+1}}{(k+1)(k+2)} - \frac{x_k}{k(k+1)} = \frac{3}{k+2} - \frac{2}{k+1}$$

Sum from $k = 1$ to $n-1$:

$$\sum_{k=1}^{n-1} \left(\frac{x_{k+1}}{(k+1)(k+2)} - \frac{x_k}{k(k+1)} \right) = \sum_{k=1}^{n-1} \left(\frac{1}{k+2} + \frac{2}{k+2} - \frac{2}{k+1} \right)$$

$$\therefore \frac{x_n}{n(n+1)} - \cancel{\frac{x_1}{2}} = \underbrace{\sum_{k=1}^{n-1} \frac{1}{k+2}}_{k \leftarrow k-2} + \left(\frac{2}{n+1} - 1 \right)$$

$$\therefore \frac{x_n}{n(n+1)} = \sum_{k=3}^{n+1} \frac{1}{k} + \left(\frac{2}{n+1} - 1 \right)$$

$$\therefore \quad \frac{X_n}{n(n+1)} = \underbrace{\sum_{k=1}^{n} \frac{1}{k}}_{H_n} + \frac{1}{n+1} - 1 - \frac{1}{2} + \frac{2}{n+1} - 1$$

$$\therefore \quad \frac{X_n}{n(n+1)} = H_n + \frac{3}{n+1} - \frac{5}{2}$$

n^{th} harmonic #

$$\therefore \quad X_n = 3n - \frac{5}{2} n(n+1) + n(n+1) H_n$$

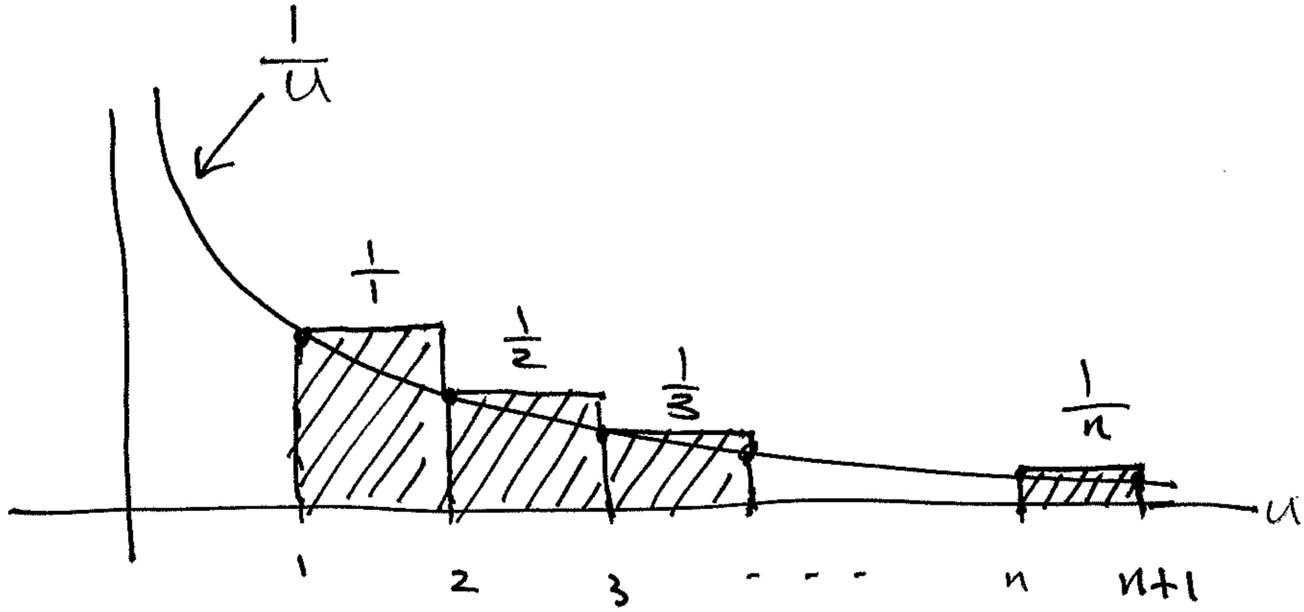
Recall

$$t(n) = (n-1) + \frac{2}{n} X_n$$

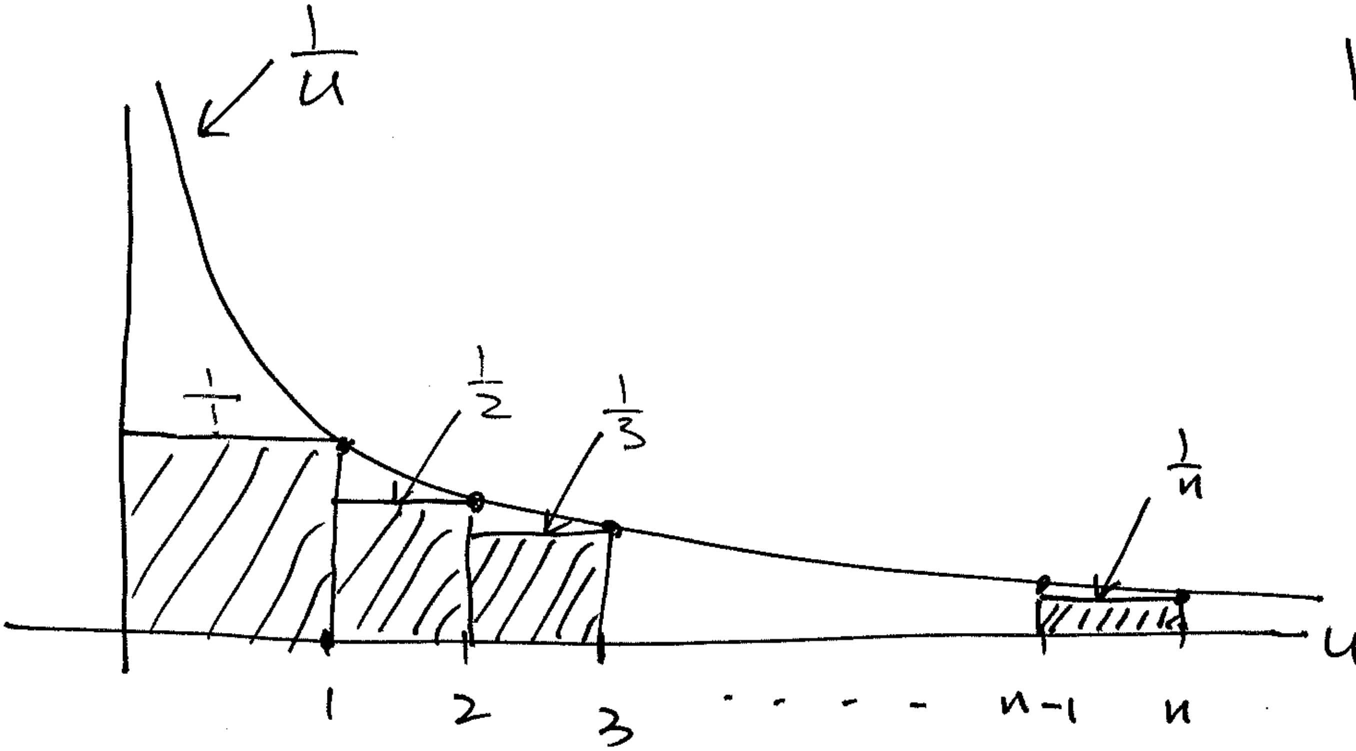
$$= (n-1) + 6 - 5(n+1) + 2(n+1) H_n$$

$$\therefore \boxed{t(n) = -4n + 2(n+1)H_n}$$

estimate size of H_n



$$\therefore \int_1^{n+1} \frac{1}{u}\,du \le H_n \le 1 + \int_1^{n} \frac{1}{u}\,du$$



$$\therefore \quad \underbrace{\ln(n+1)}_{\Omega(\ln(n))} \le H_n \le \underbrace{1+\ln(n)}_{O(\ln(n))}$$

$$\therefore \quad H_n = \Theta(\ln(n)) \quad \text{actually} \quad H_n \sim \ln(n)$$