CSE 102 2-4-20

Quicksort Sec. 7.1-7.4

sort $A[p \dots r]$ recursively

<u>Quicksort(A, p, r)</u>

		#comp
1.	if $p < r$	0
2.	q = Partition(A, p, r)	$n-1$
3.	Quicksort(A, p, q-1)	
4.	Quicksort(A, q+1, r)	

Partition causes:

$$\underbrace{A[p \dots (q-1)]}_{\text{still not sorted}} \leq \underset{\substack{\uparrow \\ \text{pivot}}}{A[q]} < \underbrace{A[(q+1) \dots r]}_{\text{still not sorted}}$$

Partition(A, p, r)

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

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

Pivot

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

$i{+}{+}$

$j{+}{+}$

if $A_j > A_r$: $j{+}{+}$

array comparison by partition

is $= n-1$ (top level)

$A[\underbrace{1 \,\ldots\ldots\, n-1}, n]$

at any level of recursion:

$A[p \,\ldots\ldots\, r-1, r]$

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

Runtime of Quicksort() depends on relative sizes of

$$A[p \cdots q-1] \ \& \ A[q+1 \cdots r]$$

worst case occurs A is already sorted

$$A[p \cdots (r-1)] \leq A[r] < \text{empty}$$

let $T(n)$ = worst case # of array comp. by Quicksort on arrays of length n.

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

Iteration:

$$T(n) = (n-1) + T(n-1)$$
$$= (n-1) + (n-2) + T(n-2)$$
$$\vdots$$
$$= \sum_{i=1}^{k} (n-i) + T(n-k)$$

stop when $n-k=1$, i.e

when $\boxed{k = n-1}$

so

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

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

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

$$= \frac{1}{2} n(n-1) = \frac{1}{2} n^2 - \frac{1}{2} n = \Theta(n^2)$$

what's Quick about this ?

what about average case ?

- Assume all Permutations of A[1 ... n] are equally likely, i.e. all have Probability $= \frac{1}{n!}$
- let $t(n)$ = average case # of array comp. by Quicksort on arrays of length n.

$$t(n) = \frac{\sum_{\text{all Permutations}} \left(\begin{array}{l} \text{\# comp. Performed by} \\ \text{Quicksort on that Perm.} \end{array} \right)}{n!}$$

<u>Goal</u>: determine a recurrence for $t(n)$

<u>note</u> our assumption implies that pivot $A[q]$ equally likely to be any of $A[1 \cdots n]$.

$$\text{Prob}(q=i) = \frac{1}{n} \quad (\text{for } 1 \le i \le n)$$

<u>note</u>:

$$\text{length}(A[1 \cdots (q-1)]) = q-1-1+1 = q-1$$

$$\text{length}(A[q+1 \cdots n]) = n-(q+1)+1 = n-q$$

Thus

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

initial value $t(0) = 0$

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

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

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

finally

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

Define

$$\begin{cases} x_n = \sum_{q=1}^{n-1} t(q) \\ x_1 = 0 \end{cases}$$

Then

$$x_{n+1} - x_n = \sum_{q=1}^{n} t(q) - \sum_{q=1}^{n-1} t(q) = t(n)$$

so

$$x_{n+1} - x_n = (n-1) + \frac{2}{n} \cdot x_n$$

$$x_{n+1} - \left(1 + \frac{2}{n}\right) x_n = (n-1)$$

$$\therefore \; x_{n+1} - \left(\frac{n+2}{n}\right) x_n = (n-1)$$

multiply by magic number: $\dfrac{1}{(n+1)(n+2)}$

$$\frac{x_{n+1}}{(n+1)(n+2)} - \frac{x_n}{n(n+1)} = \frac{(n-1)}{(n+1)(n+2)}$$

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

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 for $k=1$ to $n-1$!

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

$$\therefore \quad \frac{x_n}{n(n+1)} - \cancel{\frac{x_1}{2}} = \sum_{k=3}^{n+1}\frac{1}{k} + \frac{2}{n+1} - 1$$

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

$$= \sum_{k=1}^{n}\frac{1}{k} + \frac{3}{n+1} - \frac{5}{2}$$

Define n^{th} harmonic number

$$H_n = \sum_{k=1}^{n} \frac{1}{k}$$

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

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

Recall $\quad t(n) = (n-1) + \frac{2}{n} \cdot x_n$

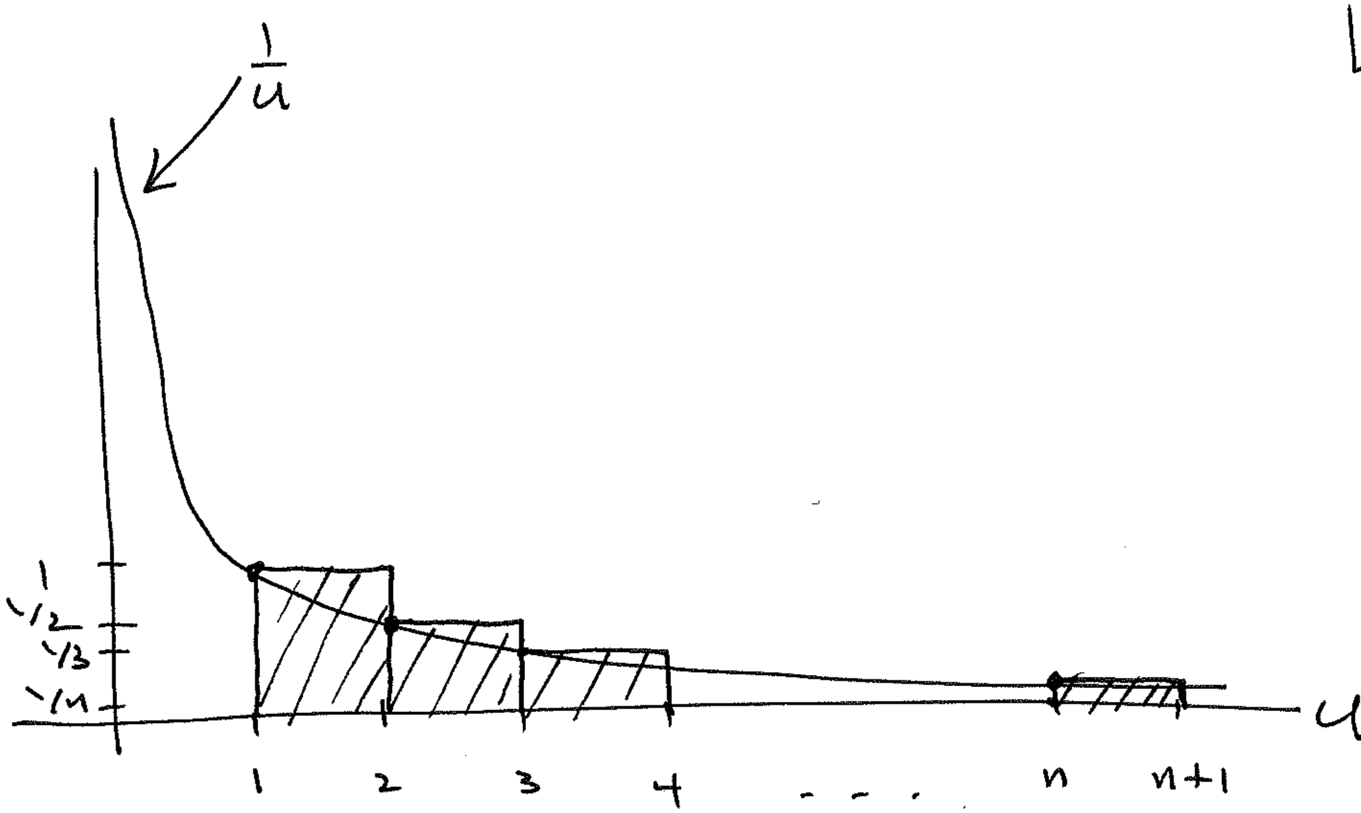
so

$$t(n) = (n-1) + \frac{2}{n}\left(3n - \frac{5}{2}n(n+1) + n(n+1)\cdot H_n\right)$$

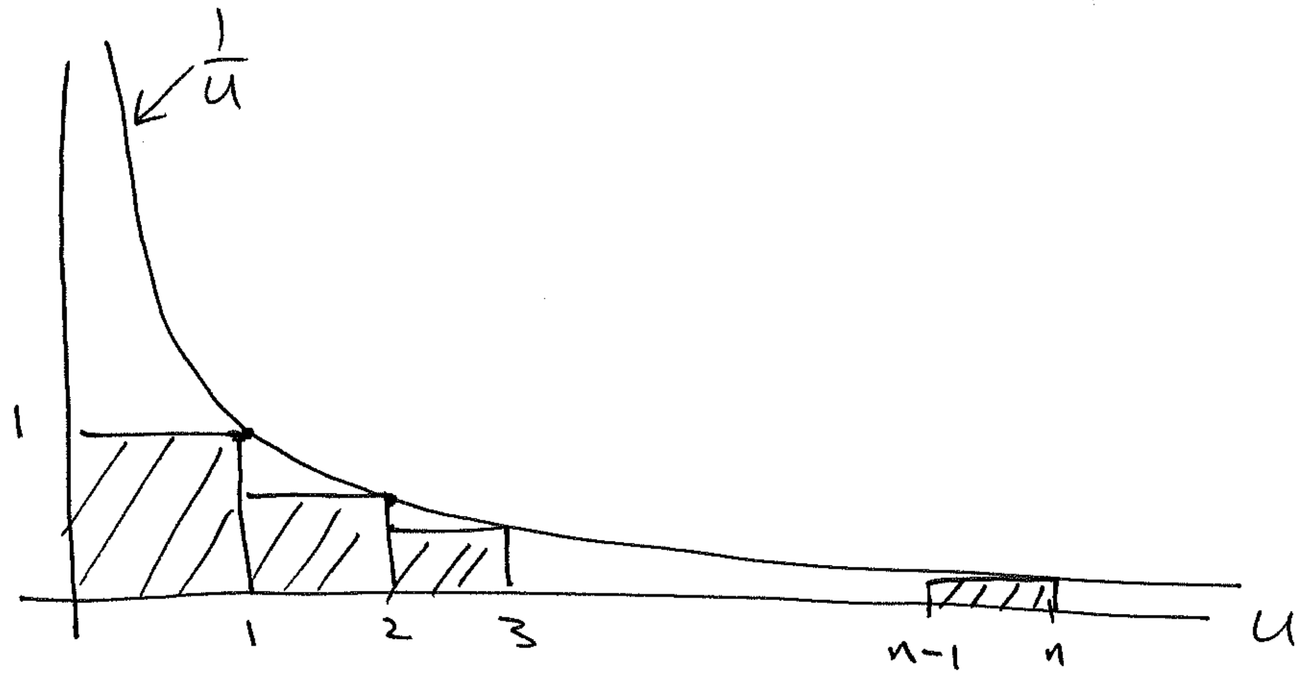
$$= n-1+6-5(n+1)+2(n+1)H_n$$

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

How ~~fast~~ does H_n grow?



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



Thus

$$\ln(u)\Big|_1^{n+1} \leq H_n \leq 1 + \ln(u)\Big|_1^{n}$$

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

$$\therefore \quad H_n = \Theta(\ln n)$$

Actually :

$$\lim_{n \to \infty} \frac{H_n}{\ln(n)} = 1$$

$$\therefore \quad H_n \sim \ln(n)$$

so

$$t(n) = 2(n+1)\cdot\Theta(\ln n) - 4n$$

$$\therefore \boxed{t(n) = \Theta(n \log n)}$$

Randomized Version:

<u>RandPartition(A, p, r)</u>

1. $i = \text{Rand}(p, r)$
2. $A[i] \leftrightarrow A[r]$ (swap)
3. return Partition(A, p, r)

RandQuicksort(A, p, r)

1. if $p < r$
2. q = RandPartition(A, p, r)
3. RandQuicksort(A, p, q-1)
4. RandQuicksort(A, q+1, r)