cse 102 11-1-22

continuing ....

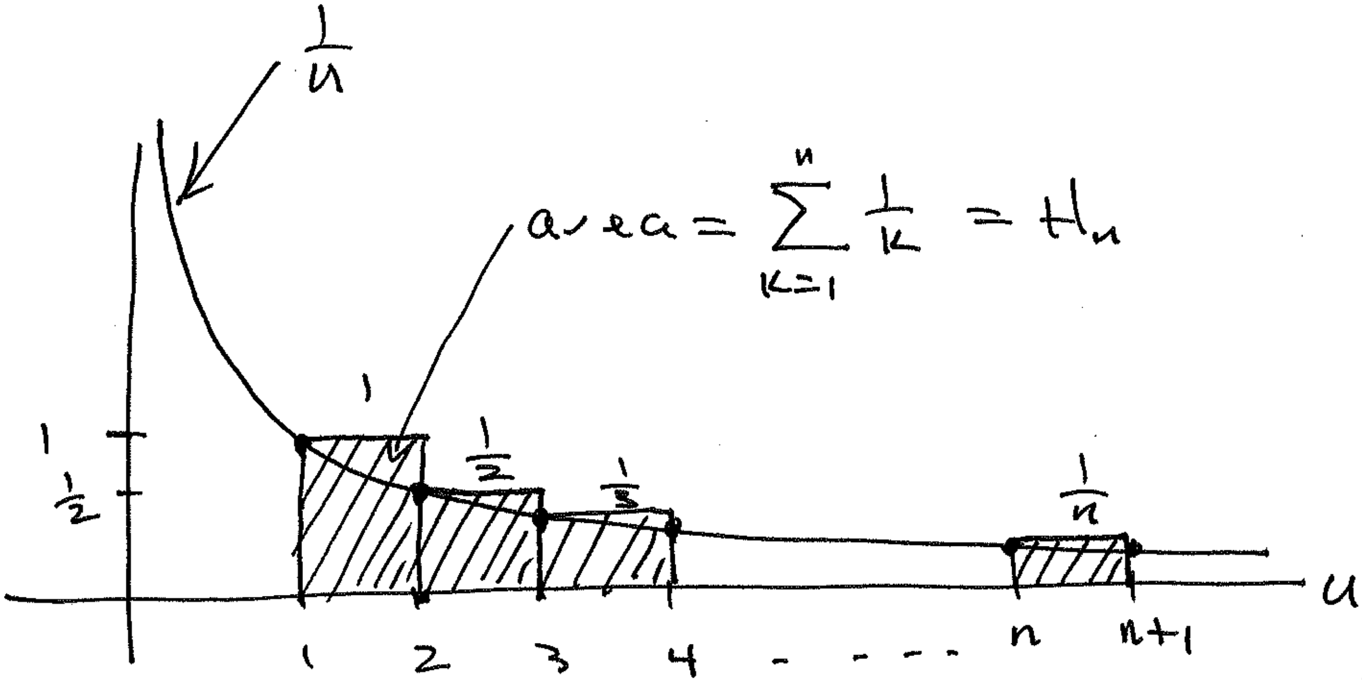
$$t(n) = (n-1) + \frac{2}{n}\left(3n - \frac{5}{2}n(n+1) + n(n+1)H_n\right)$$

$$= (n-\cancel{1}) + \cancel{6} - 5n - \cancel{5} + 2(n+1)H_n$$

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

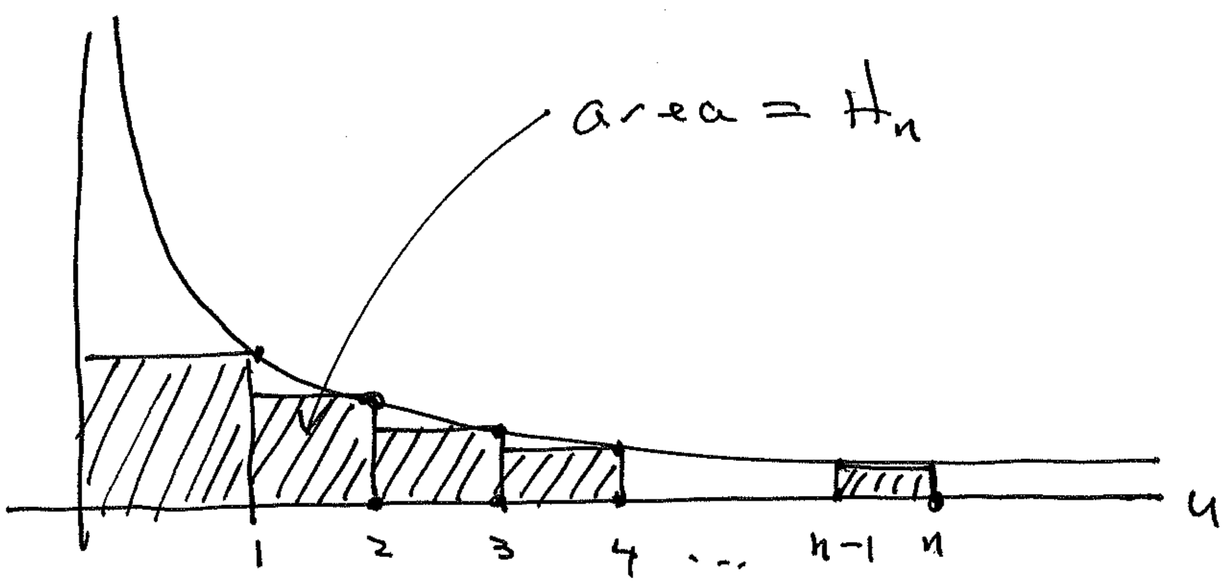
find growth rate of H_n:

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



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

Also



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

Thus

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

$$\ln(n+1) - \cancel{\ln(1)} \le H_n \le 1 + \ln(n) - \cancel{\ln(1)}$$

$$\Omega(\ln n) = \ln(n+1) \le H_n \le 1 + \ln(n) = O(\ln n)$$

$$\therefore \quad \boxed{H_n = \Theta(\ln n)}$$

Actually $H_n \sim \ln(n)$ $\left\{ \begin{array}{l} \text{Exercise!} \\ \text{Prove this.} \end{array} \right.$

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

$= 2nH_n + 2H_n - 4n$

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

Actually

$\boxed{t(n) \sim 2n\ln(n)}$

Randomization

choose a random element as Pivot in Partition.

RandPartition(A, p, r)

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

RandQuicksort(A, p, r)

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

Chap 9: Selection Problem

Ex. minMax(A,p,r) finds both min and max.

max(x, y)	#comp
return x>y ? x : y	1

min(x, y)	
return x<y ? x : y	1

minMax(A, p, r) Pre: $p \le r$	#comp
1. if p = r	0
2. return (A[p], A[p])	0
3. else	0
4. $q = \lfloor \frac{p+r}{2} \rfloor$	0
5. (m_1, M_1) = minMax(A, p, q)	$T(\lceil \frac{n}{2} \rceil)$
6. (m_2, M_2) = minMax(A, q+1, r)	$T(\lfloor \frac{n}{2} \rfloor)$
7. return (min(m_1,m_2), max(M_1,M_2))	2

let $T(n) = \#$ comp on $A[1 \cdots n]$

so

$$T(n) = \begin{cases} 0 & n = 1 \\ T\left(\lfloor \frac{n}{2} \rfloor\right) + T\left(\lceil \frac{n}{2} \rceil\right) + 2 & n \geq 2 \end{cases}$$

<u>Show</u>

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

$$\begin{aligned} RHS &= T\left(\lfloor \tfrac{n}{2} \rfloor\right) + T\left(\lceil \tfrac{n}{2} \rceil\right) + 2 \\ &= \left(2\lfloor \tfrac{n}{2} \rfloor - 2\right) + \left(2\lceil \tfrac{n}{2} \rceil - 2\right) + 2 \\ &= 2\left(\lfloor \tfrac{n}{2} \rfloor + \lceil \tfrac{n}{2} \rceil\right) - 2 - \cancel{2} + \cancel{2} \\ &= 2n - 2 = T(n) = LHS. \end{aligned}$$

$\therefore$ $T(n) = \Theta(n)$.

The selection Problem

Defn

The i^{th} order statistic of $A[1 \ldots n]$ consisting of n distinct elements is the i^{th} smallest element.

- i.e. its the element that is greater than $i-1$ other elements.
- i.e. " " " "
  " greater than or equal to i elements in $A[1 \ldots n]$.
- if $A[1 \ldots n]$ was sorted, it is the element in $A[i]$.

Problem

Given such an array, find the i^{th} order statistic.

- one way: sort first
  $\text{cost} = \Theta(n \log n)$.

- we can do better: $\underbrace{\Theta(n)}_{\text{average case}}$.

Recall: RandPartition(A, P, r) splits $A[p \ldots r]$ so that

$$A[p \ldots (q-1)] < \underset{\text{Pivot}}{A[q]} < A[(q+1) \ldots r]$$

$\downarrow$
return q

RandSelect(A, p, r, i)

Pre: $1 \le i \le r-p+1$

```
if p = r
    return A[p]
q = RandPartition(A, p, r)
k = q - p + 1        # length of A[p...q]
if k = i
    return A[q]
else if i < k
    return RandSelect(A, p, q-1, i)
else                 # i > k
    return RandSelect(A, q+1, r, i-k)
```

Exercise: Prove correctness.

let $t(n)$ = average # of array comparisons by RandSelect on $A[1 \cdots n]$

Assume all permutations of $A[1 \cdots n]$ are equally likely. Hence return value q of RandPartition$(A, 1, n)$ equally likely to be any of $1 \leq q \leq n$

Recall RandPartition$(A, 1, n)$ does $(n-1)$ comparisons

$$t(n) = \frac{\sum_{q=1}^{n} \Big((n-1) + P(i<q) \cdot t(q-1) + P(i>q) \cdot t(n-q) \Big)}{n}$$

<u>note</u>: $P(i<q) = P(i<q\leq n) = \frac{n-i}{n}$

$P(i>q) = P(1 \leq q < i) = \frac{i-1}{n}$

So

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

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

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

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

Exercise : show $t(n) = \Theta(n)$.

(see mid1 review #4 solutions)