Quicksort Average Case

Step 1. Write a formula

$$T_{AVG}(n) = \frac{1}{n}\sum_{r=1}^{n}\left\{T_{AVG}(r-1) + T_{AVG}(n-r) + O(n)\right\}$$

Step 2. Guess that $T_{AVG}(n) = O(n\log(n)) \le C\log(n)$ for some $C>0$

Step 3. Check

$$T_{AVG}(n) = \frac{1}{n}\sum_{r=1}^{n}\left\{T_{AVG}(r-1) + T_{AVG}(n-r) + O(n)\right\}$$

$$\le \frac{1}{n}\sum_{r=1}^{n}\left\{C\cdot(r-1)\log(r-1) + C\cdot(n-r)\log(n-r) + a\cdot n\right\}$$

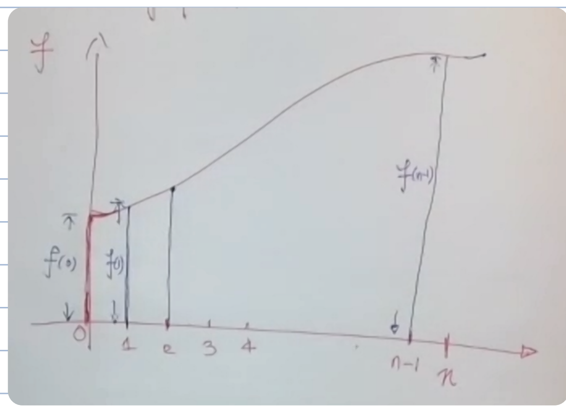
... Split the summations, pull out C

$$\frac{C}{n}\sum_{r=1}^{n}\left\{ \dots \right\}$$

$$= 2\frac{C}{n}\sum_{r=1}^{n}\left\{(r-1)(\log(r-1))\right\} + a\cdot n$$

$$f(r) = (r-1)(\log(r-1))$$

Monotonic Function



$$\le 2\frac{C}{n}\int_{0}^{n} x\log_2(x)\,dx$$

Computer science always use log base 2

$$= \frac{x^2}{2}\ln(x) - \frac{x^2}{4} + C$$

$$\int n\log_2(n)\, d \;=\; \frac{n^2}{2\ln(2)}\cdot\ln(n) - \frac{n^2}{4\ln(2)} + a\cdot n$$

$$= 2\frac{C}{n}\left(\frac{n^2}{2\ln(2)}\cdot\ln(n) - \frac{n^2}{4\ln(2)} + a\cdot n\right)$$

$$\le C\cdot n\cdot\log_2(n) \quad \text{when} \quad C >> a$$

Quick Sort Summary:

1. Worst Case $O(n^2)$, when already Sorted.
2. Average Case $O(n\log(n))$, When randomly Sorted

Must be able to do so in 7 minutes on midterm