Randomized Quicksort

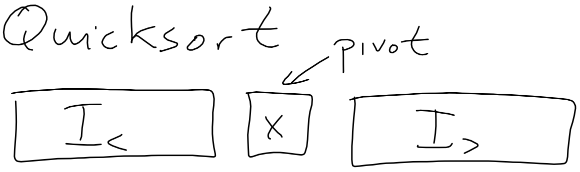
Same idea as for Randomized median but we repeat till we found a good splitter (expected # repetitions is 2)

So the expected running time on given subset S is $O(|S|)$ (excluding time for recursion)

S' is of type j if $n\left(\frac{3}{4}\right)^{j+1} \leq |S'| \leq n\left(\frac{3}{4}\right)^{j}$

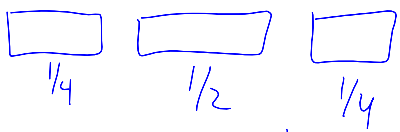
So Expected time on each problem of type j is $O\left(n\left(\frac{3}{4}\right)^{j}\right)$

In Quicksort



Recall that we repeat finding a pivot until it is a good splitter

This is expected to take 2 tries



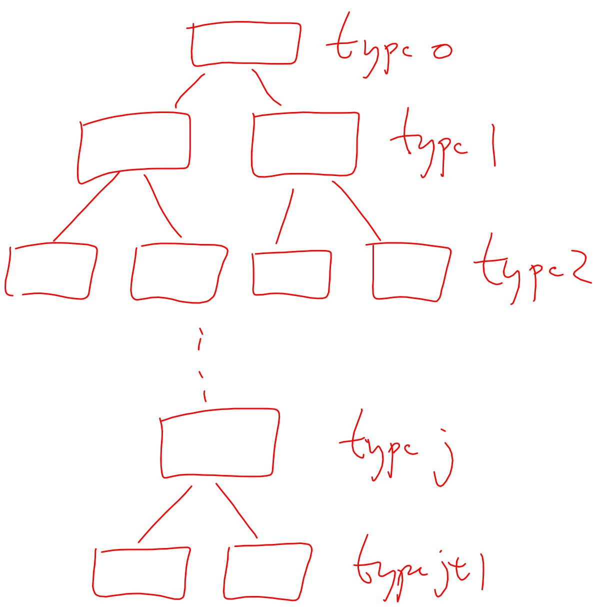
no of different types j is at most $\log_{\frac{3}{4}} n$ so $O(\log n)$

So we need to prove that Expected work on all subproblems of type j is $O(n)$

(We know that for each such S' we expect to spend $O(n(\frac{3}{4})^j)$)

Key observation:

all subproblems of type j are disjoint!



so there at most $\frac{n}{n\left(\frac{3}{4}\right)^{j+1}} \leq \left(\frac{4}{3}\right)^{j+1}$

Hence Expected work on all these

is $O\left(\left(\frac{4}{3}\right)^{j+1} n\left(\frac{3}{4}\right)^{j}\right) = O(n)$

Conclusion: Expected running time

of Rand QS is $O(n\log n)$

Analysis ala Cormen

Need to estimate # comparisons.

Let the sorted order of our set S be $s_1 \le s_2 \le \cdots \le s_n$

Define $X_{ij} = \begin{cases} 1 & \text{iff } s_i \text{ and } s_j \text{ compared} \\ 0 & \text{otherwise} \end{cases}$

Then $X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$

recall that $E(X_{ij}) = p_{ij}$: probability that s_i and s_j get compared

P_{ij} ? ..... s_i ---- s_j -----

Zoom in on interval from s_i to s_j in the sorted order:

assume $s_i, s_{i+1} \ldots s_j$ all in the same subset to be sorted

$\underbrace{\ldots\ldots}$ $s_i, s_{i+1}, \ldots, s_{j-1}, s_j$ $\underbrace{\ldots\ldots}$

such a pivot will interval unsplit

also here

Conclusion: the list $s_i, s_{i+1}, \ldots s_{j-1}, s_j$ is intact until the first pivot that lies in the list is chosen.

Case 1: first pivot is from $s_{i+1}, \ldots, s_{j-1}$
here s_i and s_j will never be compared as the will fall in different sets $s_i \in I_<$ and $s_j \in I_>$ wrt pivot

Case 2 first pivot from $s_i, s_{i+1} \ldots s_j$ is s_i or s_j

$$P_{ij} = \text{prob of case} = \frac{2}{j-i+1}$$

Now we can find $E(X)$

$$E(X) = E\left(\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} X_{ij}\right)$$

$$= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} E(X_{ij})$$

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

$$= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \frac{2}{j-i+1}$$

$$\leq \sum_{i=1}^{n-1}\sum_{r=1}^{n} \frac{1}{r} \leq n\,H(n) \sim O(n\log n)$$