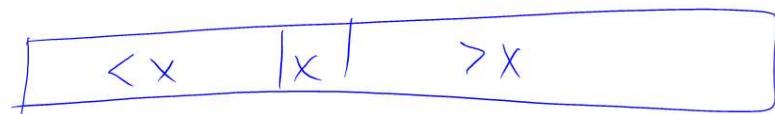


CompSci 161
Winter 2023 Lecture 05:
Divide and Conquer II:
QuickSort and Order Statistics

2 QuickSort Step 1: Partition

85	24	63	45	17	31	96	50
----	----	----	----	----	----	----	----

1. Choose a pivot.
2. Place that pivot in the right spot.
3. Pivot the rest of the array.



3

QuickSort

	↓	↓	↓			↓	↓	↓
85	24	63	45	17	31	96	50	
17	24	31	45	50	63	96	50	85

4

How fast is QuickSort?

- ▶ $T(n) = T(\text{lower}) + T(\text{upper}) + \Theta(n)$
- ▶ If lower and upper are both size $n/2$?

$$T(n) = 2T(n/2) + n = \Theta(n \lg n)$$

- ▶ What if we select a pivot uniformly at random?
- ▶ What if we could find a median in $\Theta(n)$...

5 Average Case Analysis of QuickSort

Suppose

- ▶ All permutations equally likely
- ▶ All n values are distinct (for simplicity)
- ▶ Define S_1, S_2, \dots, S_n as sorted order.

Let $P_{i,j}$ be probability we compare S_i and S_j .

$$P_{i,j} = \frac{\# \text{ yes}}{\# \text{ total}} = \frac{2}{j-i+1}$$

6 Expected number of comparisons

$$\begin{aligned} E\left(\sum_{i=1}^n \sum_{j=i+1}^n X_{i,j}\right) &= \sum_{i=1}^n \sum_{j=i+1}^n E(X_{i,j}) \\ &= \sum_{i=1}^n \sum_{j=i+1}^n \frac{2}{j-i+1} \\ &= \sum_{i=1}^n \sum_{k=2}^{n-i+1} \frac{2}{k} \\ &< \sum_{i=1}^n \left(\sum_{k=1}^n \frac{2}{k} \right) \text{ harmonic} \\ &\text{is } \Theta(n \lg n) \end{aligned}$$

7 The Selection Problem

- ▶ Given a list S and numeric k
- ▶ Want: if we sorted S , what is S_k ?
- ▶ Brute force:
 - ▶ Sort S in $\Theta(n \log n)$
 - ▶ Return S_k
- ▶ Can we do better?

8 Randomized Selection

```
quickSelect( $S, k$ )
```

If n is small, brute force and return.

Pick a random $x \in S$ and put rest into:

L , elements smaller than x

G , elements greater than x

} aka
partition

if $k \leq |L|$ **then**

 return quickSelect(L, k)

else if $k == |L| + 1$ **then**

 return x

else

 return quickSelect($G, k-|L|-1$)

9

Randomized Selection

- ▶ What is the worst-case running time?
- ▶ What would cause that bad time?
- ▶ Estimate the *expected* running time?
Hint: on average, the pivot is the median.

$$T(n) = T\left(\frac{n}{2}\right) + n$$

is $\Theta(n)$