Lecture 3.4

Yao's Lemma.

Comparison based sorting algorithm.

We know that any comparison-based deterministic sorting algorithm must make $\Omega(n \log n)$ comparisons.

Proof by decision trees.

Can randomization help us break the $\Omega(n \log n)$ lower bound?

No.

A perspective of randomized algorithms.

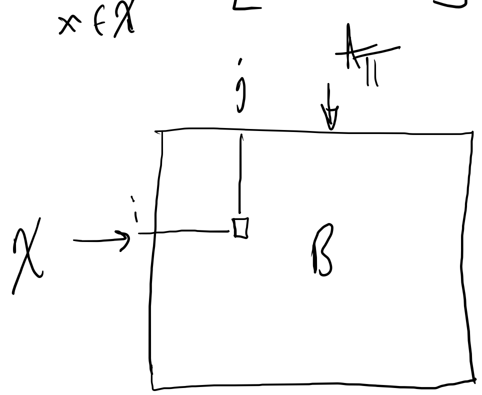
A randomized algorithm can be equivalently viewed as a probability distribution over the set $\{\underline{\underline{A}}(s_i) \mid \underline{\underline{s_i}} \in \{0,1\}^{s}\}$ of deterministic algorithms.

<u>Theorem</u> (Yao's Lemma) Let A be a <u>randomized</u> algorithm for some problem Π. For an input x, let $T(A,x)$ be the random variable denoting the "cost" of A on x. Let

$\mathcal{X}$ be the set of all inputs to Π of length n, X be a random variable having distribution p on $\mathcal{X}$, A_Π be the set of all deterministic algorithms for Π. Then

$$\max_{x \in \mathcal{X}} \mathbb{E}\left[T(A, x)\right] \geq \min_{a \in A_\Pi} \mathbb{E}\left[T(a, X)\right]$$

Proof:



$$\min_{z \in \Delta(A_\Pi)} \boxed{\max_{x \in \mathcal{X}} e_x B z} \geq \max_{y \in \Delta(\mathcal{X})} \boxed{\min_{a \in A_\Pi} y B e_a}$$

$$\Rightarrow \max_{x \in \mathcal{X}} e_x B \sigma_A \geq \min_{a \in A_\Pi} X B e_a$$

i.e. $$\max_{x \in \mathcal{X}} \mathbb{E}\left[T(A, x)\right] \geq \min_{a \in A_\Pi} \mathbb{E}\left[T(a, X)\right]$$ ∎

Theorem: Any comparison-based randomized algorithm to sort n objects must make $\Omega(n \log n)$ comparisons.

Proof: Let A be any randomized comparison-based sorting algorithm, $\mathcal{A}$ the set of all deterministic ———

" , $T(A, x)$ the random variable denoting the number of comparisons made by A on x.

To show: $\max\limits_{x \in \mathcal{X}} \mathbb{E}[T(A, x)] = \Omega(n \log n)$

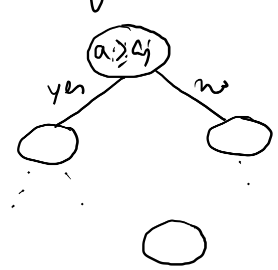
Let X be a random variable having <u>uniform distribution</u> on the set $\mathcal{X}$ of all inputs.

By Yao's lemma, it is enough to show that

$$\mathbb{E}[\,T(a, X)\,] = \Omega(n \log n) \quad \text{for every}$$

deterministic algorithm a.

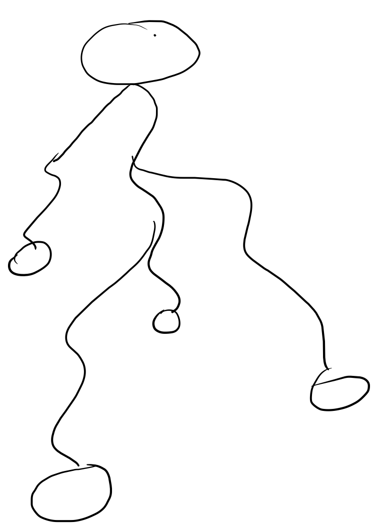
<u>To show:</u>



number of comparisons is the depth of the tree.

There are at least $n!$ leaf nodes.

The depth $\Omega(\log n!) = \Omega(n \log n)$



<u>To show:</u> Average depth of the leaf nodes of the decision tree for a is $\Omega(n \log n)$.

If the tree is balanced, then the average depth of leaf nodes is $\Omega(\log n!) = \underline{\underline{\Omega}}(n \log n)$. ∎