Formally:
Fix input size $n$.
Let $A_1, A_2, \ldots$ be all correct det algms.
Let $I_1, I_2, \ldots$ be all possible inputs.
Let $t_{ij}$ = runtime of $A_i$ on $I_j$.
Rand. algm can be viewed as picking a random $A_i$ under distribution.
Let $p_i$ = prob of picking $A_i$.
$q_j$ = prob. of picking $I_j$.

Yao's Principle (restated formally)

$\forall p_1, p_2, \ldots$, with $\sum_j p_i = 1$, $\forall q_1, q_2, \ldots$, with $\sum_i q_j = 1$

$$\max_i \sum_j p_{ij} t_{ij} \geq \min_j \sum_i q_{ij} t_{ij}$$

 expected runtime of rand. algm on $I_j$

expected runtime of $A_i$ on rand. input

Pf:
$$\max_i \sum_j p_{ij} t_{ij} \geq \sum_j q_j (\sum_i p_{ij})$$

$= \sum_{i,j} p_i q_{ij}$

$= \sum_i p_i (\sum_j q_{ij})$

$\geq \min_j \sum_i q_{ij} t_{ij}$. \qed

Remark: also follows from LP duality or von Neumann’s minimax theorem from classical 2-player game theory.

For AND-OR tree problem,
Say leaf level is AND.
Consider rand. input where at each leaf to 1 with prob. $p_i$.
Consider rand input with
we set each leaf to 1 with prob. \( p \)
0 else independently

where \( p = (1-p)^2 \)
i.e. \( p^2 - 3p + 1 = 0 \)
i.e. \( p = \frac{3 - \sqrt{5}}{2} \approx 2 - \phi \)
\( \approx 0.382 \)

Claim for AND node \( v \), \( \Pr [v \text{ has value } 1] = p \)
OR node \( v \), \( \Pr [v \text{ has value } 0] = p \)

Pf: By induction on level,

\[
\begin{align*}
\text{AND} & \quad \Pr[v \text{ is } 1] = (1-p)^2 = p \\
\text{OR} & \quad \Pr[v \text{ is } 0] = (1-p)^2 = p.
\end{align*}
\]

Consider any det alg'm on this rand input.
Let \( T_k \) = expected runtime at level \( k \).

Say OR node.

\[
\begin{align*}
\text{OR} & \quad \text{if first child examined is } 0, \\
\text{AND} & \quad \text{need 2 recursive calls.}
\end{align*}
\]

Details ... since alg'm may not be recursive ...)

\[
T_k \geq T_{k-1} + (1-p) T_{k-1} = (2-p) T_{k-1}.
\]

\[
\rightarrow T_k \geq (2-p)^k = \phi^k \quad (\phi = \frac{1+\sqrt{5}}{2})
\]

\[
\Rightarrow \text{total expected time } \Omega (e^{\log_2 N})
\]

\[
= \Omega (N^{0.694})
\]
*Random Sampling*

**Selection**

Given set $S$ of $n$ numbers and $k$,
find $k^{th}$ smallest

e.g. $k = n/2 \Rightarrow$ median

Quickselect $\Rightarrow O(n)$ expected time
Blum et al. ’73 ("median-of-medians-of-5")
$\Rightarrow O(n)$ det. time.

**Question 1:** const factor in # comps?

Blum et al. ’73 $\leq 16n$ comps
Schönhage et al. ’76 $\leq 3n$ comps
Dor, Zwick ’95 $\leq 2.95n$ comps
Current record!

**Lower bds:**

- trivial $\geq n$ comps
- Blum et al. ’73 $\geq 1.5n$ comps (by adversary)
- B. John ’80 $\geq 2n$ comps
- Dor, Zwick ’96 $\geq \frac{1}{2^{100000000000000000000000000000000000000}} n$

Can rand. help?

**Question 2:** Suppose we have limited space $s$.

(input array is read-only)

What’s the best time/space trade-off?
Floyd and Rivest's Algorithm ('75?): (Monte-Carlo version)

// to compute median

1. Pick random sample $R \leq S$ of size $r$

2. Compute $\left(\frac{r}{2} - c\sqrt{r}\right)$-th smallest $a$ of $R$
   and $\left(\frac{r}{2} + c\sqrt{r}\right)$-th smallest $b$ of $R$

3. Let $L = \{\text{all elements } \leq a\}$
   $M = \{\text{all elements between } a \cdot A \cdot b\}$
   $H = \{\text{all elements } \geq b\}$

   Check: $|L| \leq \frac{n}{2}$ and $|H| \leq \frac{n}{2}$ and $|M| \leq 4c\sqrt{\frac{n}{r}}$

4. Return $\left(\frac{n}{2} - |L|\right)$-th smallest of $M$.

Runtime: $O(r \log r) = o(n)$

Set $r = \frac{n}{3}$

Line 2: naively by sorting

$O(r \log r) = o(n)$

Line 3: $2n$ comps

~ $n + (\text{M} + \text{H})$ comps

Comp with a

Line 4: naively by sorting

$O(\frac{n}{3} \log \frac{n}{3})$

$= O(\frac{c \sqrt{n}}{\sqrt{r}} \log n) = o(n)$.
comp with a
\[ \sim \frac{3n^2}{2} + 4\frac{cn}{\sqrt{n}} \]

\[ \implies 1.5n + o(n) \text{ comps beats det.!!} \]

**Note:** 3 ways to define a rand. sample \( R \):

1. "with replacement", i.e.
   
   for \( i = 1 \) to \( r \) do
   
   pick a random element \( z_i \in S \) independently
   
   put \( z_i \) in \( R \)

   \((|R| \leq r)\)

2. "w/o replacement", i.e.
   
   for \( i = 1 \) to \( r \) do
   
   pick a random element \( z_i \in S - \{z_1, \ldots, z_{i-1}\} \) independently
   
   put \( z_i \) in \( R \)

   \((|R| = r)\)

   (each subset chosen w.

   prob \( \frac{1}{\binom{n}{r}} \))

3. flip coins, i.e.
   
   for \( z \in S \) independently decide to put \( z \) in \( R \) w. prob \( r/n \)

   \((E(|R|) = r)\)

   here, with replacement. (option 1)

**Error Analysis:**

\[ \Pr[\text{error}] = \Pr(\left(\ast\right) \text{ false}) \]
\[ 1M \leq c \sqrt{n} \]

\[
\leq \Pr \left( \text{ranks}(a) < \frac{n}{2} - 2c \frac{n}{\sqrt{v}} \right.
\text{ or } \text{ranks}(a) > \frac{n}{2} \\
\text{ or } \text{ranks}(b) < \frac{n}{2} \\
\text{ or } \text{ranks}(b) > \frac{n}{2} + 2c \frac{n}{\sqrt{v}} \right)
\]

\[
\Pr \left( \text{ranks}(a) < \frac{n}{2} - 2c \frac{n}{\sqrt{v}} \right) \leq ?
\]

TO BE CONT'D...