

**Goal:** to derive

$$\begin{cases} \mathbb{E}(X_n) & := \# \text{ probes per } \mathbf{succ} \text{ search in } n\text{-node BST} \\ \mathbb{E}(Y_n) & := \# \text{ probes per } \mathbf{unsucc} \text{ search} \quad " \end{cases}$$

**Sample Spaces:**

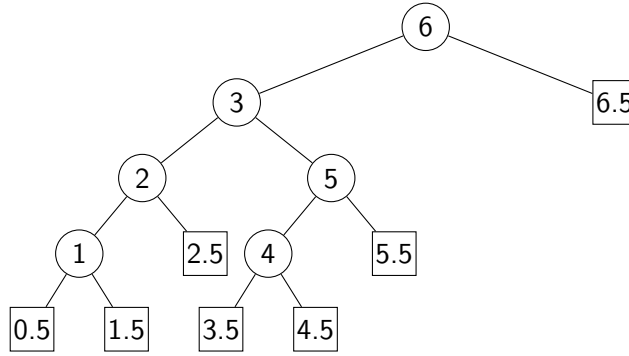
$$\begin{cases} \Omega_{X_n} & = \left\{ (x_1, x_2, \dots, x_n; k) \mid \begin{array}{l} (x_1, x_2, \dots, x_n) \in S_n \\ 1 \leq k \leq n \end{array} \right\}, \quad |\Omega_{X_n}| = n!n \\ \Omega_{Y_n} & = \left\{ (x_1, x_2, \dots, x_n; y) \mid \begin{array}{l} (x_1, x_2, \dots, x_n) \in S_n \\ y = 0.5, 1.5, \dots, n.5 \end{array} \right\}, \quad |\Omega_{Y_n}| = (n+1)! \end{cases}$$

$S_n$  = set of all permutations (relative order of keys) of  $\{1, 2, \dots, n\}$   
 $\Omega_{Y_n} \approx \{(x_1, x_2, \dots, x_n; y) \mid (x_1, x_2, \dots, x_n; y) \in S_{n+1}\}$ , by re-numbering

**Random Variables:**

$$\begin{cases} X_n((x_1, x_2, \dots, x_n; k)) & := \# \text{ probes to } \mathbf{succ} \text{ search } x_k \text{ in } T(x_1, x_2, \dots, x_n) \\ Y_n((x_1, x_2, \dots, x_n; y)) & := \# \text{ probes to } \mathbf{unsucc} \text{ search } y \quad " \end{cases}$$

$T(x_1, x_2, \dots, x_n)$  := BST formed by inserting keys  $x_1, x_2, \dots, x_n$   
 $T(6, 3, 2, 1, 5, 4) =$



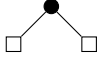
$$\begin{cases} \sum_{1 \leq k \leq 6} X_6((6, 3, 2, 1, 5, 4; k)) & = (1 + 2 + 3 + 4 + 3 + 4) = 17 = 6 + I(T(6, 3, 2, 1, 5, 4)) \\ \sum_{y=0.5, \dots, 6.5} Y_6((6, 3, 2, 1, 5, 4; y)) & = (4 + 4 + 3 + 4 + 4 + 3 + 1) = 23 = E(T(\cdot)) \end{cases}$$

$$\begin{cases} I(T) & := \sum_{x: \text{int node of } T} \text{path length}(x) = 11 \\ E(T) & := \sum_{y: \text{ext node of } T} \text{path length}(y) = 23 = I(T) + 2 \cdot 6 \end{cases}$$

$$\mathbb{E}(X_6) = \frac{1}{6} \left[ (1 + \mathbb{E}(Y_0)) + (1 + \mathbb{E}(Y_1)) + \dots + (1 + \mathbb{E}(Y_4)) + (1 + \mathbb{E}(Y_5)) \right]$$

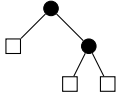
n	1	2	3	
$\mathbb{E}(X_n) = 2 \frac{n+1}{n} H_n - 3$	1	$\frac{3}{2}$	$\frac{17}{9}$	$\dots$
$\mathbb{E}(Y_n) = 2H_{n+1} - 2$	1	$\frac{5}{3}$	$\frac{13}{6}$	$\dots$

$n = 1$

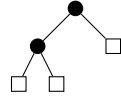


$$\begin{cases} \mathbb{E}(X_1) = \frac{1}{1} [1] = 1 \\ \mathbb{E}(Y_1) = \frac{1}{2} [1 + 1] = 1 \end{cases}$$

$n = 2$



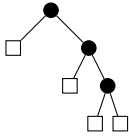
(1, 2)



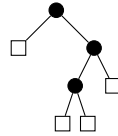
(2, 1)

$$\begin{cases} \mathbb{E}(X_2) = \frac{1}{2 \cdot 2!} [(1 + 2) + (1 + 2)] = \frac{3}{2} \\ \mathbb{E}(Y_2) = \frac{1}{3 \cdot 2!} [(1 + 2 + 2) + (2 + 2 + 1)] = \frac{5}{3} \end{cases}$$

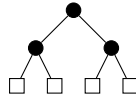
$n = 3$



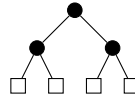
(1, 2, 3)



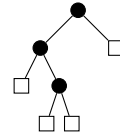
(1, 3, 2)



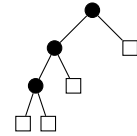
(2, 1, 3)



(2, 3, 1)



(3, 1, 2)



(3, 2, 1)

$$\begin{cases} \mathbb{E}(X_3) = \frac{1}{3 \cdot 3!} [(1 + 2 + 3) + (1 + 2 + 3) + (1 + 2 + 2) + (\cdot) + (\cdot) + (\cdot)] = \frac{17}{9} \\ \mathbb{E}(Y_3) = \frac{1}{4 \cdot 3!} [(1 + 2 + 3 + 3) + (1 + 3 + 3 + 2) + (2 + 2 + 2 + 2) + (\cdot) + (\cdot) + (\cdot)] = \frac{13}{6} \end{cases}$$

**Remark:**

$\begin{cases} \text{Dynamic model} & : n! \text{ trees;} \\ \text{Static} & " & : \frac{1}{n+1} \binom{2n}{n} \text{ trees, } [T(2, 1, 3) = T(2, 3, 1)] \end{cases}$

**Lemma 1.**  $n$ -node binary tree  $T$  has

- (1)  $2n$  edges,
- (2)  $n + 1$  external nodes (leaves),
- (3)  $E(T) = I(T) + 2n$ .

**Lemma 2.**  $(n + 1)\mathbb{E}(Y_n) = n\mathbb{E}(X_n) + n$

**Proof:**

$$\begin{aligned}
(1) \mathbb{E}(X_n) &= \sum_{w \in \Omega_{X_n}} P(w) X_n(w) = \sum_{\substack{(x_1, \dots, x_n) \in S_n \\ 1 \leq k \leq n}} \frac{1}{n! n} X_n(x_1, \dots, x_n; k) \\
&= \frac{1}{n! n} \sum_{(x_1, \dots, x_n) \in S_n} \sum_{1 \leq k \leq n} X_n(x_1, \dots, x_n; k) \\
&= \frac{1}{n! n} \sum_T (n + I(T)) \\
&= \frac{1}{n} \left[ n + \frac{1}{n!} \sum_T I(T) \right] \\
(2) \mathbb{E}(Y_n) &= \frac{1}{(n + 1)!} \sum_{(x_1, \dots, x_n) \in S_n} \sum_{y=0.5, \dots, n.5} Y_n(x_1, \dots, x_n; y) \\
&= \frac{1}{(n + 1)!} \sum_T E(T) \\
&= \frac{1}{(n + 1)n!} \sum_T (2n + I(T)) \\
&= \frac{1}{n + 1} \left[ 2n + \frac{1}{n!} \sum_T I(T) \right]
\end{aligned}$$

$$(1) + (2) \Rightarrow (n + 1)\mathbb{E}(Y_n) = n\mathbb{E}(X_n) + n \quad \blacksquare$$

**Lemma 3.**  $\mathbb{E}(X_n) = 1 + \frac{1}{n} \sum_{1 \leq k \leq n} \mathbb{E}(Y_{k-1})$

**Theorem.**

$$\begin{cases} \mathbb{E}(X_n) &= 2 \frac{n+1}{n} H_n - 3 \approx 2 \ln n \\ \mathbb{E}(Y_n) &= 2H_{n+1} - 2 \approx 2 \ln n \end{cases}$$

**Proof:** Lemmas 2 and 3 give

$$\begin{aligned}
(n + 1)\mathbb{E}(Y_n) &= \sum_{1 \leq k \leq n} \mathbb{E}(Y_{k-1}) + 2n \\
- ) \quad n\mathbb{E}(Y_{n-1}) &= \sum_{1 \leq k \leq n-1} \mathbb{E}(Y_{k-1}) + 2(n - 1)
\end{aligned}$$

---


$$(n + 1)\mathbb{E}(Y_n) = (n + 1)\mathbb{E}(Y_{n-1}) + 2$$

$$\Rightarrow \begin{cases} \mathbb{E}(Y_n) &= \mathbb{E}(Y_{n-1}) + \frac{2}{n+1} \\ \mathbb{E}(Y_{n-1}) &= \mathbb{E}(Y_{n-2}) + \frac{2}{n} \\ &\vdots \\ \mathbb{E}(Y_2) &= \mathbb{E}(Y_1) + \frac{2}{3} \\ \mathbb{E}(Y_1) &= \frac{2}{2} \end{cases}$$

$$\Rightarrow \begin{cases} \mathbb{E}(Y_n) &= 2H_{n+1} - 2 \\ \mathbb{E}(X_n) &= \frac{n+1}{n} E(Y_n) - 1 = 2\frac{n+1}{n} H_n - 3 \end{cases} \quad \blacksquare$$

**Lemma 3.**  $\mathbb{E}(X_n) = 1 + \frac{1}{n} \sum_{1 \leq k \leq n} \mathbb{E}(Y_{k-1})$

**Proof:**

$$\begin{aligned} \mathbb{E}(X_n) &= \frac{1}{n!n} \sum_{(x_1, \dots, x_n)} \sum_{1 \leq k \leq n} X_n(x_1, \dots, x_n; k) \\ &= \frac{1}{n!n} \sum_{(x_1, \dots, x_n) \in \mathcal{S}_n} \sum_{1 \leq k \leq n} \left(1 + Y_n(x_1, \dots, x_{k-1}; k)\right) \\ &= 1 + \frac{1}{n!n} \sum_{1 \leq k \leq n} \sum_{(x_1, \dots, x_k) \in \mathcal{S}_k} n \cdots (k+1) Y_{k-1}(x_1, \dots, x_{k-1}; x_k) \\ &= 1 + \frac{1}{n} \sum_{1 \leq k \leq n} \left[ \frac{1}{k!} \sum_{(x_1, \dots, x_k) \in \mathcal{S}_k} Y_{k-1}(x_1, \dots, x_{k-1}; x_k) \right] \\ &= 1 + \frac{1}{n} \sum_{1 \leq k \leq n} \mathbb{E}(Y_{k-1}) \quad \blacksquare \end{aligned}$$