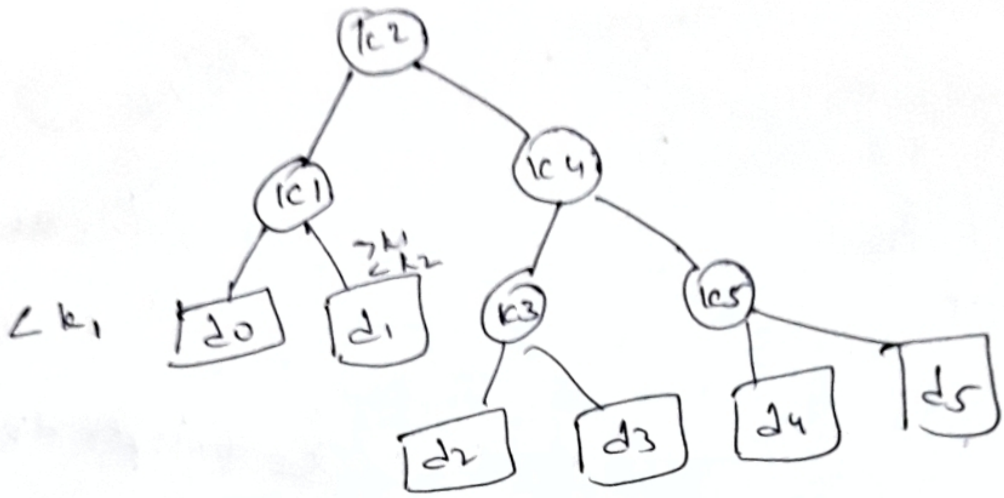
OBST

Suppose we have a tree



a search can be successful or unsuccessful.
Unsuccessful reach dummy keys.
successful end at any keys.

d_i refers k_i and k_{i+1}

p_i probability of successful key k_i

probability for dummy key

q_i probability of d_i

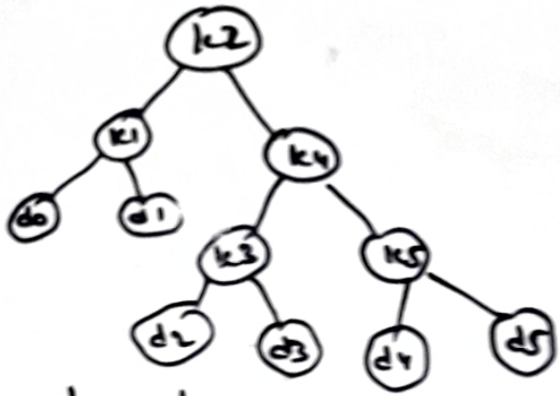
Since probability is always equal to 1 for whole search

$$\Rightarrow \sum_{i=1}^{n} p_i + \sum_{i=0}^{n} q_i = 1 \longrightarrow (1)$$

sum of probabilities

$k1 < k2 < k3 < k4 < k5$

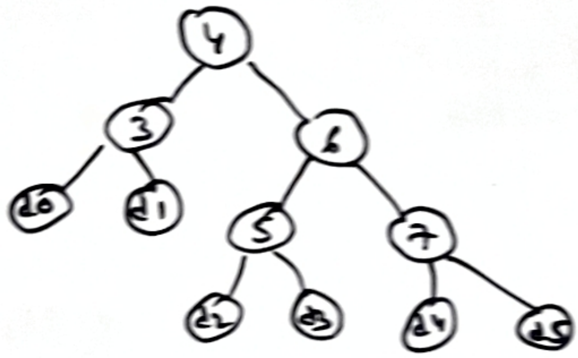
$d0 < d_1 < d_2 < d3 < d4 < d5$



$d_0 \ldots d_5$ dummy nodes

total dummy nodes $= \left[\text{no. of internal nodes}\right] + 1$

for ex :-



Search 4.5 (Not there) reaches $\underline{d2}$ — probability q_2 — non occurrence of d2.

∴ probability of occurrences of elements, k_i is P_i

Probability of non-occurrence of elements, d_i is q_i

so total probability of finding a particular element ~~is~~ in OBST $\sum_{i=1}^{n} P_i$

and not finding will be $\sum_{i=0}^{n} q_i$

$$\therefore \sum_{i=1}^{n} p_i + \sum_{i=0}^{n} q_i = 1$$

Expected cost of a search in T is

$$= \sum_{i=1}^{n} (dept(k_i) + 1) p_i + \sum_{i=0}^{n} (dept(d_i)+1) q_i$$

$$= \sum_{i=1}^{n} dept(k_i) . p_i + \sum_{i=1}^{n} p_i + \sum_{i=0}^{n} dept(d_i) q_i + \sum_{i=0}^{n} q_i$$

$$= \left[\sum_{i=1}^{n} p_i + \sum_{i=0}^{n} q_i \right] + \sum_{i=1}^{n} dept(k_i) . p_i + \sum_{i=0}^{n} dept(d_i) q_i$$

$$= 1 + \sum_{i=1}^{n} dept(k_i) p_i + \sum_{i=0}^{n} dept(d_i) q_i$$

But this cost varies if rooted with different k_i, different OBST.

using DP

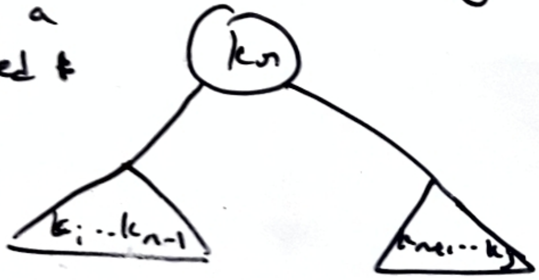
Step 1:- Optimal Substructure

Consider subtree of BST $k_i, k_{i+1}, \dots k_j$, $1 \le i \le j \le n$

so the leafs or dummy nodes $d_{i-1}, d_i, d_{i+1} \dots d_j$

lets say a tree rooted with k_r

$i \le r \le j$



Total probability

$$w(i,j) = p_r + w(i, r-1) + w(r+1, j)$$

Step 2 :- Recursive Solution

In $k_i \ldots k_j$, when $j = i-1$, there are no ~~active~~ actual keys, only dummy keys d_{i-1}

Let $e[i,j]$ denote expected cost of searching an

$$e[i, i-1] = q_{i-1}$$

Let us denote

$$w(i,j) = \sum_{l=i}^{j} p_l + \sum_{l=i-1}^{j} q_l$$

be the probability $k_i \ldots k_j$

If k_r is root in $k_i \ldots k_j$ then

$$c[i,j] = \underbrace{p_r}_{\text{probability of root}} + (\underbrace{c[i, r-1]}_{\text{cost of left subtree}} + \underbrace{w[i, r-1]}_{\text{probability of left subtree}}) + (c(r+1, j) + w(r+1, j)] \rightarrow (1)$$

right right

and

$$w[i,j] = w[i, r-1] + p_r + w[r+1, j] \rightarrow (2)$$

from (1) & (2)

$$c[i,j] = \begin{cases} q_{i-1} & \text{if } j = i-1 \\ \min\limits_{i \le r \le j} \{ c[i, r-1] + c[r+1, j] + w[i,j) & \text{if } i \le j \end{cases}$$

Instead of computing $w(i,j)$ every time, we can also write it as

$$w(i,j) = w(i, j-1) + p_j + q_j$$

Initially $w(i,i) = q(i)$, $c(i,i) = 0$, $k(i,i) = 0$

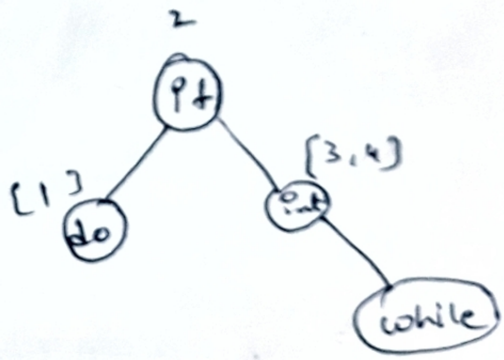
Construct OBST for the instance of $n = 4$,

$(a_1, a_2, a_3, a_4) = \{$ do, if, int, while $\}$.

$p(1:4) = (3, 3, 1, 1)$

$q(0:4) = (2, 3, 1, 1, 1)$

	0	1	2	3	4
0	$w_{00} = 2$ $c_{00} = 0$ $r_{00} = 0$	$w_{11} = 3$ $c_{11} = 0$ $r_{11} = 0$	$w_{22} = 1$ $c_{22} = 0$ $r_{22} = 0$	$w_{33} = 1$ $c_{33} = 0$ $r_{33} = 0$	$w_{44} = 1$ $c_{44} = 0$ $r_{44} = 0$
1	$w_{01} = 8$ $c_{01} = 8$ $r_{01} = 1$	$w_{12} = 7$ $c_{12} = 7$ $r_{12} = 2$	$w_{23} = 3$ $c_{23} = 3$ $r_{23} = 3$	$w_{34} = 3$ $c_{34} = 3$ $r_{34} = 4$	
2	$w(0,2) = 12$ $c_{02} = 19$ $r_{02} = 1$	$w_{13} = 9$ $c_{13} = 12$ $r_{13} = 2$	$w_{24} = 5$ $c_{24} = 8$ $r_{24} = 3$		
3	$w_{03} = 14$ $c_{03} = 25$ $r_{03} = 2$	$w_{14} = 11$ $c_{14} = 19$ $r_{14} = 2$			
4	$w_{04} = 16$ $c_{04} = 32$ $r_{04} = 2$				



Initially $w(i,i) = q(i)$

$c(i,i) = 0$

$r(i,i) = 0$

$\Rightarrow$ $w_{00} = q(0) = 2$

$w_{11} = q(1) = 3$

$w_{22} = q(2) = 1$

$w_{33} = q(3) = 1$

$w_{44} = q(4) = 1$

$w(i,j) = p(j) + q(j) + w(i, j-1)$

$c(i,j) = \min\limits_{i<k\le j} \{ c(i, k-1) + c(k, j) \} + w(i,j)$

$w(0,1) = p(1) + q(1) + w(0,0)$

$= 3 + 3 + 2$

$= 8$

$c(0,1) = \min\limits_{0\le k\le 1} \{ c(0,0) + c(1,1) \} + w(0,1)$

$= 8$

$r(0,1) = 1$

$c(1,2) = \min\limits_{1<k\le 2} \{ c(1,1) + c(2,2) \} + w(1,2)$

$w(1,2) = p(2) + q(2) + w(1,1)$

$= 3 + 1 + 3 = 7$

$c(1,2) = 0 + 0 + 7 = 7$

$r(1,2) = 2$

$w(2,3) = p(3) + q(3) + w(2,2)$

$= 1 + 1 + 1 = 3$

$c(2,3) = 0 + 0 + 3 = 3$

$r(2,3) = 3$

$w(3,4) = p(4) + q(4) + w(3,3)$

$= 1 + 1 + 1 = 3$

$c(3,4) = \min \{ c(3,3) + c(4,4) \} + w(3,4)$

$= 0 + 0 + 3 = 3$

$r(3,4) = 4$

$$w(0,2) = p(2) + q(2) + w(0,1)$$
$$= 3 + 1 + 8 = 12$$

$$c(0,2) = \min_{0<k\le 2} \begin{cases} k=1 & c(0,0) + c(1,2) + w(0,2) \\ k=2 & c(0,1) + c(2,2) + w(0,2) \end{cases}$$

$$= \min \begin{cases} k=1 & 0+7+12 = 19 \\ k=2 & 8+0+12 = 20 \end{cases}$$

$$c(0,2) = 19$$
$$r(0,2) = k = 1$$

$$w(1,3) = p(3) + q(3) + w(1,2)$$
$$= 1 + 1 + 7 = 9$$

$$c(1,3) = \min_{0\le k\le 3} \begin{cases} k=2 & c(1,1) + c(2,3) + w(1,3) \\ k=3 & c(1,2) + c(3,3) + w(1,3) \end{cases}$$

$$= \min \begin{cases} k=2 & 0+3+9 = 12 \\ k=3 & 7+0+9 = 16 \end{cases}$$

$$c(1,3) = 12$$
$$r(1,3) = 2$$

$$w(2,4) = 5, \quad c_{24} = 8, \quad r_{24} = 3$$

$$w(0,3) = p(3) + q(3) + w(0,2)$$
$$= 1 + 1 + 12 = 14$$

$$c(0,3) = \min_{0<k\le 3} \begin{cases} k=1 & c(0,0) + c(1,3) + w(0,3) \\ k=2 & c(0,1) + c(2,3) + w(0,3) \\ k=3 & c(0,2) + c(3,3) + w(0,3) \end{cases}$$

$$= \min \begin{cases} k=1 & 0+12+14 = 26 \\ k=2 & 8+3+14 = 25 \\ k=3 & 19+0+14 = 33 \end{cases}$$

$$c(0,3) = 25$$
$$r(0,3) = 2$$