

Assignment 1

Shreekara Raghavan - EE18BTECH11040

1 PROBLEM

Let T be a full binary tree with 8 leaves. (A full binary tree has every level full.) Suppose two leaves a and b of T are chosen uniformly and independently at random. The expected value of the distance between a and b in T (i.e., the number of edges in the unique path between a and b) is (rounded off to 2 decimal places)

2 FULL BINARY TREE

- A full binary tree is a binary tree whose nodes (apart from the leaf nodes) have 2 children.
- If a binary tree has leaf nodes in the form of 2^k , then these kinds of binary trees becomes a perfect binary tree whose height is k .

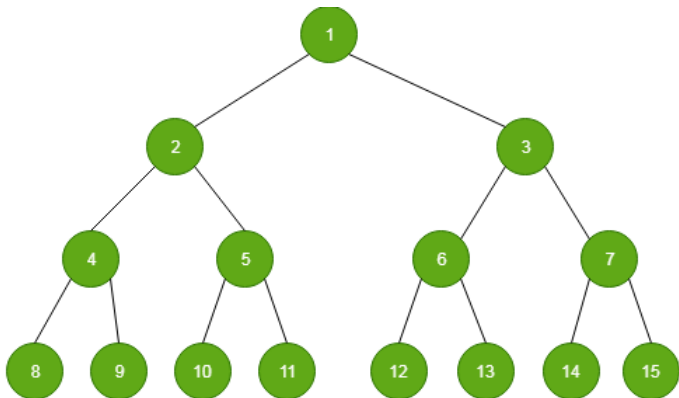


Fig. 0: Full binary Tree with 8 Leaf Nodes

3 DISTANCE BETWEEN 2 LEAF NODES

There exists a unique path between 2 nodes of a binary tree without retracing the same edge. The length of the path taken depends on which leaf nodes we choose, specifically at what level their common ancestor is above the leaf nodes.

- If the common ancestor of the leaf nodes (say a and b) is placed 1 level above the leaf node, (the 2 leaf nodes are sibling nodes)

then the path length would be 1 (distance between node a and the parent node) + 1 (distance between parent node and node b) = 2.

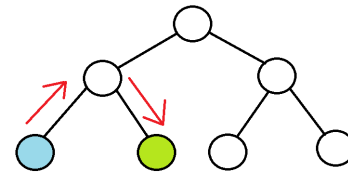


Fig. 0: distance when common ancestor is 1 level above = 2

- Similarly, if the common ancestor of the leaf nodes is placed 2 levels above the leaf node, then the path length would be 2 (distance between node a and the common ancestor) + 2 (distance between common ancestor and node b) = 4.

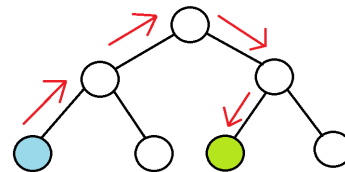


Fig. 0: distance when common ancestor is 2 levels above = 4

In short, we can conclude,

$$d = 2 * l \quad (3.0.1)$$

where d is the distance between the 2 nodes
 l is the number of levels between the leaf nodes and their common ancestor.

3.1 Implementation

Below is the code for calculating distance between 2 given nodes.

https://github.com/Shreekara277/Data_Structures/tree/main/codes

4 EXPECTED DISTANCE

Expected distance is given by

$$d = \sum_n x \cdot p_X(x) \quad (4.0.1)$$

We can select 2 leaf nodes among 8(repetitions included) in 8×8 ways = 64 ways. Thus, $n = 64$. Since the nodes are chosen uniformly at random, probability of choosing a pair is $\frac{1}{64}$. Thus the equation modifies to:

$$d = \frac{1}{64} \sum_n x \quad (4.0.2)$$

The summation represents the sum of distances between all possible combinations of leaf nodes. If the pair of leaf nodes contain the same node, then the distance is taken to be 0.

4.1 Solution

Below is the code for calculating expected distance.

https://github.com/Shreekara277/Data_Structures/tree/main/codes

We get the total distance as 272. Expected length = 4.25