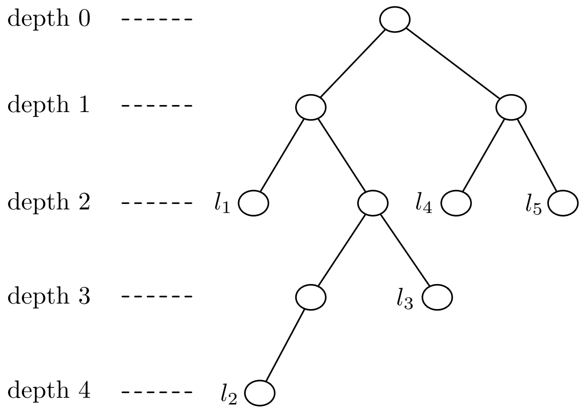
▶ **Problem DS-04-07** Suppose a binary tree has leaves $l_1, l_2, \ldots, l_M$ at depths $d_1, d_2, \ldots, d_M$, respectively. Prove that $\sum_{i=1}^{M} s^{-d_i} \leq 1$ and determine when the equality is true.

Proof. Before the proof, we first give an example to see the inequality $\sum_{i=1}^{M} s^{-d_i} \leq 1$. Consider a tree with five leaves (i.e., $M = 5$) shown below.



Then, $d_1 = depth(l_1) = 2$, $d_2 = depth(l_2) = 4$, $d_3 = depth(l_3) = 3$, $d_4 = depth(l_4) = 2$, and $d_5 = depth(l_5) = 2$. Thus,

$$\sum_{i=1}^{5} 2^{-d_i} = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^3} + \frac{1}{2^2} + \frac{1}{2^2} = \frac{4+1+2+4+4}{16} = \frac{15}{16} \leq 1$$

For any tree T with M leaves, we denote by $F(T) = \sum_{i=1}^{M} s^{-d_i}$ where $d_i = depth(l_i)$ for each leaf l_i in T. The proof is by induction on the number of nodes in a tree. Clearly, in a tree T with no nodes, we have $F(T) = 0$, and in a one-node tree T, the root is a leaf at depth zero, so $F(T) = 1$ and the claim is true. Suppose the result is true for all trees with at most N nodes. Consider any tree T with $N+1$ nodes. Without loss of generality, we assume that such a tree consists of a left subtree T_L with k node and a right subtree T_R with $N-k$ node. By the inductive hypothesis, $F(T_L) \leq 1$ and $F(T_R) \leq 1$. Because all leaves are one deeper with respect to the original tree T than with respect to the subtrees T_L or T_R, $F(T) = \frac{1}{2}F(T_L) + \frac{1}{2}F(T_R) \leq 1$, proving the result.

The equality is true if and only if there are no nodes with one child. If there is a node with one child, the equality cannot be true because adding the second child would increase the sum to higher than 1. If no nodes have one child, then we can find and remove two sibling leaves, creating a new tree. It is easy to see that this new tree has the same sum as the old. Applying this step repeatedly, we arrive at a single node, whose sum is 1. Thus the original tree had sum 1.

$\square$