Homework 5: Exercise 6.42

Maryam Bahrani (mbahrani)

Dylan Mavrides

Let $\mathcal{T}$ be the class of binary trees.

For a random binary tree t of size N, let $c_0(t)$ be the number of internal nodes of t with *no* external children, $c_1(t)$ be the average number of internal nodes with exactly *one* external children, and $c_2(t)$ the average number of internal nodes with *two* external children.

First, we will compute the average value of c_2 over random binary trees using the CGF method from lecture:

$$C_2(z) = \sum_{t \in \mathcal{T}} c_2(t) z^{|t|}$$

Furthermore, the number of internal nodes with two external children is the sum of the number internal nodes with two external children in the root's left subtree and in the root's right subtree, giving the following functional equation for $c_2(t)$ when $t = z \times t_L \times t_R$:

$$c_2(t) = c_2(t_L) + c_2(t_R)$$

We can use this too simplify the CGF $C_2(z)$

well-argued

$$\begin{aligned}
C_2(z) &= \sum_{t \in \mathcal{T}} c_2(t) z^{|t|} \\
&= \sum_{t \in \mathcal{T}, |t| <= 1} c_2(t) z^{|t|} + \sum_{t \in \mathcal{T}, |t| > 1} c_2(t) z^{|t|} \\
&= 0 + 1 \cdot z + \sum_{t_L \in \mathcal{T}} \sum_{t_R \in \mathcal{T}} (c_2(t_L) + c_2(t_R)) z^{|t_L| + |t_R| + 1} \\
&= z + z \sum_{t_L \in \mathcal{T}} \sum_{t_R \in \mathcal{T}} (c_2(t_L) + c_2(t_R)) z^{|t_L| + |t_R|} \\
&= z + 2z C_2(z) T(z),
\end{aligned}$$

where the last step is identical to what was used in slide 36 of lecture 6. Solving for $C_2(z)$, we have

$$C_2(z) = \frac{z}{1 - 2zT(z)} = \frac{z}{\sqrt{1 - 4z}}.$$

Finally, the radius-of-convergence transfer theorem with $\rho = 1/4$, $\alpha = 1/2$, and $f(z) = z$ (whose radius of convergence is indeed infinite and hence greater than ρ) gives

$$[z^N]C_2(z) \sim \frac{4^{N-1}}{\sqrt{N\pi}}.$$

The average number of internal nodes with two external children is therefore

$$\frac{[z^N]C_2(z)}{[z^N]T(z)} \sim \frac{\frac{4^{N-1}}{\sqrt{N\pi}}}{\frac{4^N}{\sqrt{N^3\pi}}} = \frac{N}{4}.$$

Denote by X_i the expectation of c_i for trees of size N, $i = 0, 1, 2$. First, note that the number of external nodes in a binary tree is equal to twice the number of internal nodes with two external children plus the number of internal nodes with one external child. Since there are $N + 1$ external nodes in a tree of size N, by linearity of expectation,

$$\begin{aligned} 2X_2 + X_1 &= N + 1 \\ X_1 &\sim \frac{N}{2} + 1 \end{aligned}$$

good

Finally, since $X_0 + X_1 + X_2 = N$, we must have

$$X_0 = N - X_1 - X_2 \sim N - (\frac{N}{2} + 1) - \frac{N}{4} = \frac{N}{4} - 1.$$

Overall, the fraction of internal nodes with two external children is $\sim 1/4$, the fraction of internal nodes with exactly one external child is $\sim 1/2$, and the fraction of internal nodes with no external children is $\sim 1/4$.