The height of a random leaf in conditioned Galton-Watson tree

Douglas Rizzolo
Rough Notes

1 Leaves in Binary Trees

Let T_n be a uniformly random rooted ordered binary tree with n leaves (equivalently, $2n-1$ verticles) and, conditionally given T_n, let L_n be a uniformly random vertex of T_n. This lecture is devoted to the answering the question, what is the height of U_n?

Our goal is to prove the following result:

Theorem 1. *If $(x_n)_{n\geq 0}$ is a sequence of integers such that $x_n/\sqrt{n} \to x \in (0,\infty)$ then*

$$\lim_{n\to\infty} \sqrt{n}\mathbb{P}(ht(L_n) = k_n) = \frac{x}{2}\exp\left(-\frac{x^2}{4}\right).$$

This implies that

$$\frac{1}{\sqrt{n}} ht(L_n) \xrightarrow{d} \text{Rayleigh}\left(\sqrt{2}\right).$$

There are many proofs of this, but we will use a generating function argument.

Proof. Let B_n be the number of binary trees with n leaves and let

$$B(z) = \sum_{n=1}^{\infty} B_n z^n.$$

Decomposing at the root, we see that for $n \geq 1$, C_n satisfies the relationship

$$B_n = \sum_{k=1}^{n-1} B_k B_{n-k}.$$

Thus we identify $B_n = C_{n-1}$, where C_n is the n'th Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Let $\mathcal{T}_n$ be the set of rooted ordered binary tree with n leaves and let $\mathcal{T} = \cup_n \mathcal{T}_n$. Note that

$$B(z) = \sum_{\mathbf{t}\in\mathcal{T}} z^{|\mathbf{t}|}$$

where $|\mathbf{t}|$ is the number of leaves of $\mathbf{t}$. Let $\theta_k(\mathbf{t})$ be the number of vertices of $\mathbf{t}$ at height k and let

$$\Theta_k(z) = \sum_{\mathbf{t}\in\mathcal{T}} \theta_k(\mathbf{t}) z^{|\mathbf{t}|}.$$

Observe that

$$\sqrt{n}\mathbb{P}(\mathrm{ht}(L_n) = k_n) = \sqrt{n} \sum_{\mathbf{t}\in\mathcal{T}_n} \frac{\theta_k(\mathbf{t})}{n} \frac{1}{B_n} = n^{-1/2}\mathbb{E}(\theta_k(T_n)) = \frac{n^{-1/2}[z^n]\Theta_k(z)}{[z^n]B(z)}.$$

Let $\mathbf{t}_l$ and $\mathbf{t}_r$ be the left and right subtrees attached to the root of $\mathbf{t}$. The for $k \geq 1$ we have

$$\begin{aligned}
\Theta_k(z) = \sum_{\mathbf{t}\in\mathcal{T}} \theta_k(\mathbf{t}) z^{|\mathbf{t}|} &= \sum_{\mathbf{t}\in\mathcal{T}} (\theta_{k-1}(\mathbf{t}_l) + \theta_{k-1}(\mathbf{t}_r)) z^{|\mathbf{t}_l|+|\mathbf{t}_r|} \\
&= \sum_{(\mathbf{t}_1,\mathbf{t}_2)\in\mathcal{T}^2} (\theta_{k-1}(\mathbf{t}_1) + \theta_{k-1}(\mathbf{t}_2)) z^{|\mathbf{t}_1|+|\mathbf{t}_2|} \\
&= \sum_{(\mathbf{t}_1,\mathbf{t}_2)\in\mathcal{T}^2} \theta_{k-1}(\mathbf{t}_1) z^{|\mathbf{t}_1|+|\mathbf{t}_2|} + \sum_{(\mathbf{t}_1,\mathbf{t}_2)\in\mathcal{T}^2} \theta_{k-1}(\mathbf{t}_2) z^{|\mathbf{t}_1|+|\mathbf{t}_2|} \\
&= 2B(z)\Theta_{k-1}(z).
\end{aligned}$$

Continuing inductively, and using the fact that $\Theta_0(z) = z$ we find that

$$\Theta_k(z) = z(2B(z))^k = 2^k z B(z)^k.$$

Thus we must analyze $[z^n]B(z)^k$, the coefficient of z^n in $B(z)^k$. Note that

$$B(z)^k = \sum_{n=0}^{\infty} \sum_{\substack{(n_1,\ldots,n_k) \\ n_1+\cdots+n_k=n}} B_{n_1} \cdots B_{n_k} z^n.$$

Equivalently,

$$[z^n]B(z)^k = \sum_{\substack{(n_1,\ldots,n_k) \\ n_1+\cdots+n_k=n}} B_{n_1} \cdots B_{n_k},$$

is the number of ordered forests of k rooted ordered binary trees with a total of n leaves among the k trees. To work with this, we use a different interpretation of the Catalan numbers. In particular, we use the fact that $B_n = C_{n-1}$ is the number of rooted ordered trees with n vertices. There is a very nice combinatorial proof of this relationship. The map in Figure 1 gives a bijection from rooted ordered binary trees with n leaves to rooted ordered trees with n vertices.

Consequently, $[z^n]B(z)^k$ is also the number of forests of k rooted ordered trees with a total of n vertices among the k trees. Thus, if we take $\hat{T}_1, \ldots, \hat{T}_k$ to be Galton-Watson trees with Geometric(1/2)-offspring distribution and let $\#\mathbf{t}$ be the number of vertices in the tree $\mathbf{t}$, we have

$$\frac{1}{2^{2n-k}}[z^n]B(z)^k = \mathbb{P}\left(\#\hat{T}_1 + \cdots + \#\hat{T}_k = n\right).$$

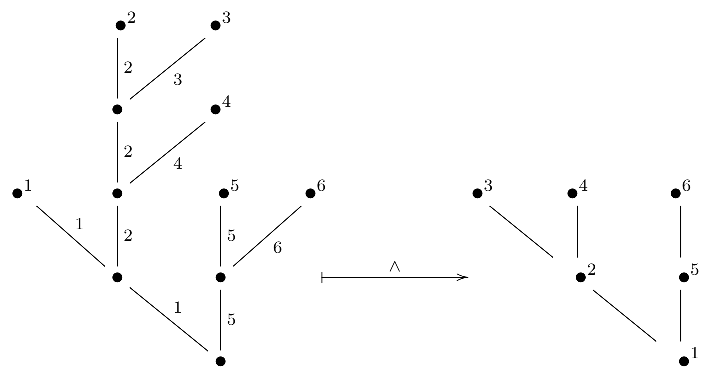

Figure 1: A binary tree $\mathbf{t}$ and its image under $\mathbf{t} \mapsto \hat{\mathbf{t}}$

Let $X_1, X_2, \dots$ be i.i.d. Geometric(1/2) random variables and let $S_n = X_1 + \cdots + X_n$. By the Otter-Dwass formula we have

$$\mathbb{P}\left(\#\hat{T}_1 + \cdots + \#\hat{T}_k = n\right) = \frac{k}{n}\mathbb{P}(S_n - n = -k) = \frac{k}{n}\binom{2n-k-1}{n-k}\frac{1}{2^n}.$$

Exercise 1. *It follows from the above that* $[z^n]B(z)^k = \frac{k}{n}\binom{2n-k-1}{n-k}$. *Give a direct combinatorial proof of this.*

Since $\mathrm{var}(X_1) = 2$ and the distribution of $X_1 - 1$ is aperiodic, the Local Central Limit Theorem implies that

$$\sup_{k \in \mathbb{Z}} \left| \sqrt{2n}\mathbb{P}(S_n - n = k) - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{k^2}{4}\right) \right| \underset{n \to \infty}{\longrightarrow} 0.$$

Consequently, if $k_n/\sqrt{n} \to x \in (0, \infty)$, we have

$$\sqrt{2n}\mathbb{P}(S_n - n = k_n) \sim \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{k_n^2}{4n}\right) \sim \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{4}\right).$$

Exercise 2. *Prove the above relationship directly from Stirling's formula instead of using the local central limit theorem.*

Consequently, we have

$$[z^n]\Theta_{k_n}(z) = 2^{k_n}[z^{n-1}]B(z)^{k_n} \sim 2^{2(n-1)}\frac{x}{2\sqrt{\pi}n} \exp\left(-\frac{x^2}{4}\right).$$

Using Stirling's formula, or the $k = 1$ case of the above argument, we see that

$$[z^n]B(z) \sim 2^{2(n-1)}\frac{1}{n^{3/2}\sqrt{\pi}}.$$

Combining all of the above, we see that

$$\sqrt{n}\mathbb{P}(\mathrm{ht}(L_n) = k_n) = \frac{n^{-1/2}[z^n]\Theta_k(z)}{[z^n]B(z)} \sim \frac{2^{2(n-1)}\frac{x}{2\sqrt{\pi}n^{3/2}}\exp\left(-\frac{x^2}{4}\right)}{2^{2(n-1)}\frac{1}{n^{3/2}\sqrt{\pi}}} = \frac{x}{2}\exp\left(-\frac{x^2}{4}\right),$$

as desired. $\square$

Exercise 3. *Let W_n be a uniformly random vertex chosen from T_n. Using a similar argument as above, find*

$$\lim_{n\to\infty} \sqrt{n}\mathbb{P}(ht(W_n) = k_n),$$

when $k_n/\sqrt{n} \to x \in (0,\infty)$.

2 Beyond Binary Trees

A modification of the above approach works in a more general case. Let μ be a critical offspring distribution with finite non-zero variance σ^2, and let T be a μ-Galton-Watson tree. Suppose that for sufficiently large n, $\mathbb{P}(|T| = n) > 0$, and for all such n let T_n be distributed like T conditioned to have n leaves. We let

$$\phi(z) = \sum_{n=0}^{\infty} \mu(n) z^n$$

and

$$B(z) = \sum_{\mathbf{t}\in\mathcal{T}} \mathbb{P}(T = \mathbf{t}) z^{|t|},$$

so that $[z^n]B(z) = \mathbb{P}(|T| = n)$. In this case, we find that

$$\Theta_k(z) = z(\phi'(B(z)))^k.$$

One may then argue as above, but this is slightly complicated by the fact that we don't have a replacement for $\hat{T}$. That is, we need to find an offspring distribution ν such that if $\hat{T}$ is a ν-Galton-Watson tree then $\mathbb{P}(|T| = n) = \mathbb{P}(\#\hat{T} = n)$. This can be done and in fact, recalling the map $\wedge$ above, $\hat{T}$ is a Galton Watson tree with offspring distribution ν, which is the distribution of

$$Y = 1 + \sum_{i=1}^{\inf\{i:X_i=0\}} (X_i - 1),$$

where $X_1, X_2, \ldots$ are i.i.d with distribution μ. Moreover, from Wald's equations, give $\mathbb{E}(Y) = 1$ and $\mathrm{var}(Y) = \sigma^2 \mathbb{E}\inf\{i : X_i = 0\} = \sigma^2/\mu(0)$.

Moreover, we must interpret $\phi'(B(z))$ probabilistically. In the binary case, $\phi(z) = \frac{1}{2} + \frac{1}{2}z^2$, so that $\phi'(z) = 1$, and there is nothing to do. However, the situation is more complicated in the general case. In the general case, we have that

$$\phi'(z) = \sum_{n=1}^{\infty} n\mu(n) z^{n-1}.$$

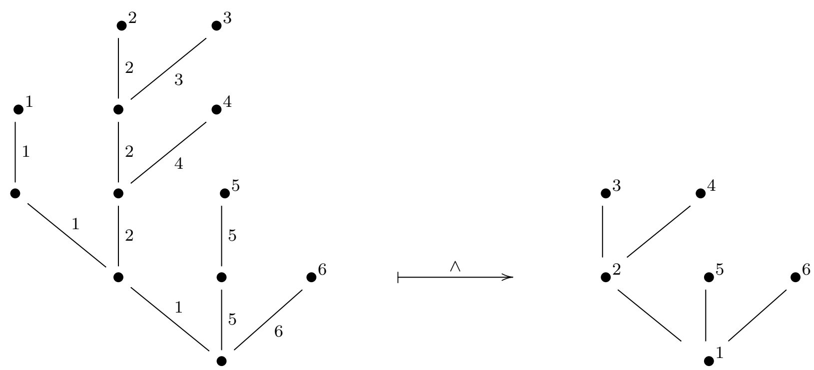

Figure 2: A tree $\mathbf{t}$ and its image under $\mathbf{t} \mapsto \hat{\mathbf{t}}$

Since μ has mean 1, we see that $\phi'(1) = 1$. Taking $\bar{\mu}(n) = n\mu(n)$, we have that $\bar{\mu}$ is a probability distribution. It is called the size-biased distribution of μ. Let $\hat{T}_1, \hat{T}_2, \ldots$ be i.i.d ν-Galton-Watson trees and let N be an independent random variable with distribution $\hat{\mu}$. Letting

$$Z = \sum_{i=1}^{N} \#\hat{T}_i,$$

we find that

$$\gamma(z) = \sum_{n=1}^{\infty} \mathbb{P}(Z = n) z^n = \phi'(B(z)).$$

Thus, if $Z_1, Z_2, \ldots$ are i.i.d with distribution Z then

$$[z^n]\phi'(B(z))^k = \mathbb{P}(Z_1 + \cdots + Z_k = n).$$

But, if $N_1, N_2, \ldots$ are i.i.d distributed like N, independent also of the $\hat{T}_i$, then

$$\mathbb{P}\left(Z_1 + \cdots + Z_k = n\right) = \mathbb{P}\left(\sum_{i=1}^{N_1 + \cdots + N_k} \#\hat{T}_i = n\right).$$

We are now left with applying a random index version of the Otter-Dwass formula, followed by a random index version of the local limit theorem. The only issues that arise in this application are resolved by using the Law of Large Numbers, which yields

$$\frac{1}{n} \sum_{j=1}^{n} N_j \xrightarrow[n \to \infty]{a.s.} \mathbb{E} N_1 = 1 + \sigma^2.$$

Exercise 4. *Fill in the details of, and complete, the above argument.*