Tree Asymptotics

Clayton Ristow

July 9 2015

Now that we have established recursive formulas for trees, we wish to develop asymptotic forumla for the coefficients of our generating functions. An expression g_n is said to be asymptotic to a sequence a_n if

$$\frac{g_n}{a_n} \to 1 \text{ as } n \to \infty \tag{1}$$

In order to get at these asymptotic formulas, we will use the following theorem to rewrite our generating function and to generate approximate forms of the coefficients of the generating function:

Theorem 1. *Let $H(x, y(x))$ be a complex vauled function that is analytic in a neighborhood of $(x_0, G(x_0))$. If the following conditions are met:*

1. $H(x_0, G(x_0)) = 0$

2. $G(x)$ *is analytic for $|x| < |x_0|$ where x_0 is the unique singularity of $G(x)$*

3. $G(x_0) = \sum_{n=0}^{\infty} G_n x_0^n$

4. $H(x, G(x)) = 0$ *if $|x| < |x_0|$*

5. $\left.\frac{\partial H}{\partial y(x)}\right|_{(x_0, G(x_0))} = 0$

6. $\left.\frac{\partial^2 H}{\partial y(x)^2}\right|_{(x_0, G(x_0))} \neq 0$

Then

$$G(x) = G(x_0) + \sum_{k=1}^{\infty} a_k (x_0 - x)^{\frac{k}{2}}$$

and if $a_1 \neq 0$ then

$$G_n \to \sqrt{\frac{a_1^2 x_0}{4\pi}} x_0^{-n} n^{\frac{-3}{2}}$$

or if $a_1 = 0$ and $a_3 \neq 0$ then

$$G_n \to \sqrt{\frac{9a_3^2 x_0^3}{16\pi}} x_0^{-n} n^{\frac{-5}{2}}$$

To use this theorem we will think of our generating function as a function of complex numbers rather than meerly a power series to generate coefficients. Also note that conditions 2,3 require us to find an x_0 such that for any number greater than x_0 $G(x)$ does not converge and then for any number less than or equal to x_0 the series does converge. In other words, x_0 is the radius f convergence for our generating function $G(x)$

1 Planted Trees

Let us begin by applying the above theorem to our functional equation for planted trees. But first we must show that all conditions are met. We must first define:

$$H(x, y) \equiv x^2 + xy + \frac{y^2}{2} + \frac{F(x^2)}{2} - y \tag{2}$$

Clearly from our functional equation $y = F(x)$ is a unique solution to $H = 0$. But notice that $H(x, F(x))$ is only zero if $F(x)$ converges. Thus if we assume for the time being that there exists an x such that

$F(x)$ actually does converge (we will show this is true). Then conditions 1 and 4 hold. Additionally, we can deduce that $\partial_y H = 0$. There is a theorem called the Implicit Function Theorem that states that if $\partial_y H \neq 0$ then there exists a function $y(x)$ that is a unique solution to $H = 0$ and $y(x_0)$ is analytic. But we know this to not be true because $F(x)$ is a unique solution to $H = 0$ and $F(x)$ is not analyitc at x_0 (different from convergent) so by contradiction we can conclude

$$\frac{\partial H}{\partial y} = 0 \text{ when } y = F(x) \text{ and } x = x_0 \tag{3}$$

So condition 5 holds. Additionally, we can simply differentiate Equation 2 with respect to y twice to get

$$\frac{\partial^2 H}{\partial y^2} = 1 \tag{4}$$

So condition 6 holds. All that remains to show are conditions 2 and 3.

We will first show that there are x's for which $F(x)$ converges. We start by recalling the recursive relation for F_n:

$$F_n = F_{n-1} + \frac{1}{2}\sum_{i=2}^{n-2} F_i F_{n-i} + \frac{1}{2}F_{\frac{n}{2}} \tag{5}$$

We can adjust the relation by changing F_1 to 1 rather than 0. Note that when this change is implemented, none of the other terms in the sqeuence will change. We may now include the F_{n-1} in the sum to get:

$$F_n = \frac{1}{2}\sum_{i=1}^{n-1} F_i F_{n-i} + \frac{1}{2}F_{\frac{n}{2}} \tag{6}$$

Now we note that F_n is an increasing series. We can obvioulsy write:

$$F_{\frac{n}{2}} < F_n$$

which then implies that

$$\frac{1}{2}F_{\frac{n}{2}} < F_n - \frac{1}{2}F_{\frac{n}{2}}$$

and then we may write

$$\frac{1}{2}F_{\frac{n}{2}} < \frac{1}{2}\sum_{i=1}^{n-1} F_i F_{n-i}$$

and we get that:

$$F_n < \sum_{i=1}^{n-1} F_i F_{n-i}$$

Since F_n is an increasing sequence we can write

$$F_n \leq \sum_{i=1}^{n-1} F_i F_{n-1-i} \tag{7}$$

The sequence of numbers

$$C_n = \sum_{i=1}^{n-1} C_i C_{n-1-i} \quad \text{where } C_1 = 1 \tag{8}$$

is called the Catalan Numbers. Thier generating function $C(x)$ has a known radius of convergence of $\frac{1}{4}$. Since $F_n \leq C_n$, we may then conclude by the Comparison Test that $F(x)$ converges for all $x \leq \frac{1}{4}$. This means that the radius of convergence of $F(x)$, x_0 must be

$$\frac{1}{4} \leq x_0$$

Additionally, we know that in order for $F(x_0)$ to have any chance of converging $F_n x_0^n$ must go to zero so at the very least $x_0 < 1$. So we may then put x_0 in the interval:

$$\frac{1}{4} \leq x_0 < 1 \tag{9}$$

Then we can say that A is the set of numbers, x, such that $F(x)$ converges. We have shown from Equation 9 that A is bounded above by 1 and nonempty so $Sup(A)$ exists. We can then write:

$$x_0 \equiv Sup(A) \tag{10}$$

We have shown that x_0, the radius of convergence of $F(x)$ exists. Now we will show how to compute it. By the Ratio Test the following is true for all x:

- if $\frac{F_{n+1}x^{n+1}}{F_n x^n} \to r < 1$ then $\sum_{n=0}^{\infty} F_n x^n$ converges

- if $\frac{F_{n+1}x^{n+1}}{F_n x^n} \to r > 1$ then $\sum_{n=0}^{\infty} F_n x^n$ diverges

So clearly the radius of convergence must be x such that $\frac{F_{n+1}x^{n+1}}{F_n x^n} \to 1$. Thus,

$$\frac{F_{n+1}x_0}{F_n} \to 1$$

and then we may write

$$x_0 = \lim_{n \to \infty} \frac{F_n}{F_{n+1}} \tag{11}$$

Finally we must show that $\lim_{x \to x_0^-} F(x)$ exists. To do this we must simply show that $F(x)$ is bounded above for all $x < x_0$ since $F(x)$ is clearly a mnotonically increasing function. To do this let us look at the functional equation:

$$F(x) = x^2 + xF(x) + \frac{1}{2}F^2(x) + \frac{1}{2}F(x^2) \tag{12}$$

Let us assume $0 < x < x_0$. Then since $F_n \geq 0$ for all n, $x^2 > 0$, $xF(x) > 0$, and $\frac{1}{2}F(x^2) > 0$. So we may then simply write

$$F(x) > \frac{1}{2}F^2(x) \tag{13}$$

And then it follows that

$$F(x) < 2 \tag{14}$$

Thus F(x) is bounded above by 2 so $\lim_{x \to x_0^-} F(x)$ exists. It follows then that $F(x_0)$ exists. Thus we may finally conclude that conditions 2 and 3 hold and we may finally apply Theorem 1 to F(x).

Before we derive the asymptotic formula, it is convenient to find the value of $F(x_0)$. To do this we first differentiate Eq. 2 with respect to y:

$$\frac{\partial H}{\partial y} = x + y - 1 \tag{15}$$

and then we plug in our solution $y = F(x)$ and evaluate at x_0 to get:

$$\left.\frac{\partial H}{\partial y}\right|_{(x_0, F(x_0))} = x_0 + F(x_0) - 1 \tag{16}$$

and then if we remember condition 5 we can write:

$$F(x_0) = 1 - x_0 \tag{17}$$

We are finally ready to expand F(x) using Theorem 1:

$$F(x) = F(x_0) + \sum_{n=1}^{\infty} a_k (x_0 - x)^{\frac{k}{2}} \tag{18}$$

and differentiaing gives us

$$F'(x) = -\sum_{n=1}^{\infty} \frac{k}{2} a_k (x_0 - x)^{\frac{k}{2}-1} \tag{19}$$

we can combine the two in the following convenient way

$$F'(F(x_0) - F(x)) = \frac{a_1^2}{2} + \mathcal{O}(x_0 - x) \tag{20}$$

So we can write

$$\lim_{x \to x_0} F'(F(x_0) - F(x)) = \frac{a_1^2}{2}$$

But we also have another way of computing ths limit using our functional equation for $F(x)$. If we differentiate this expression with respect to x we get

$$F'(x) = 2x + F(x) + xF'(x) + F(x)F'(x) + xF'(x^2) \tag{21}$$

which can be rearranged as

$$F'(x) - F(x)F'(x) = 2x + F(x) + xF'(x) + xF'(x^2) \tag{22}$$

and then we can subtract $x_0 F$ from both sides to get

$$F'(x)(F(x_0) - F(x)) = 2x + F(x) + xF'(x) - x_0 F(x) + xF'(x^2)$$

and we obtain the limit as

$$\lim_{x \to x_0} F'(F(x_0) - F(x)) = 2x_0 + F(x_0) + x_0 F'(x_0^2)$$

which simplfies to

$$\lim_{x \to x_0} F'(F(x_0) - F(x)) = x_0(1 + F'(x_0^2)) + 1 \tag{23}$$

Then combining equations 12 and 23 and solving for a_1 we get

$$a_1 = \sqrt{2(x_0(1 + F'(x_0^2)) + 1)} \tag{24}$$

And we can obtain our final asyptotic formula by plugging a_1 into Theorem 1 and simplfying:

$$F_n \approx \sqrt{\frac{x_0^2(1 + F'(x_0^2)) + x_0}{2\pi}} x_0^{-n} n^{-3/2} \tag{25}$$

x_0 and $F(x_0^2)$ can be approximated by calculating them with the power series for $F(x)$ out to as many terms as we want. When we do that we can plug the results into Eq. 25 to get

$$F_n \approx .3187766(.4026975)^{-n} n^{-3/2} \tag{26}$$

2   Rooted Trees

We will now apply the same techniques to approximate rooted trees. Before we apply Theorem 1, we must first confirm that all 6 conditions are met. Recall the functional equation for rooted trees:

$$R(x) = \frac{F(x)}{x} + \frac{F^3(x)}{6x^3} + \frac{F(x^2)F(x)}{2x^2} + \frac{F(x^3)}{3x^2} \tag{27}$$

First let us redefine $H(x, y)$ for rooted trees.

$$H(x, y) = \frac{F(x)}{x} + \frac{F^3(x)}{6x^3} + \frac{F(x^2)F(x)}{2x^2} + \frac{F(x^3)}{3x^2} - y \tag{28}$$

We can see that $y = R(x)$ is a unique solution and $R(x)$ is not analytic at its radius of convergence (We will show this value is x_0). So by the Implicit Function Theorem, condition 5 holds (just as it did for planted trees). Similarly, condition 6 holds as well. Then we can show that the radius of convergence of $R(x)$ is x_0. $R(x)$ clearly converges if $F(x)$, $F(x^2)$, and $F(x^3)$ all converge. From the previous section on planted trees we know that $F(x)$ converges for $x \le x_0$, $F(x^2)$ converges for $x \le \sqrt{x_0}$, and $F(x^3)$ converges for $x \le \sqrt[3]{x_0}$. Since $x_0 < 1$, $x_0 < \sqrt{x_0} < \sqrt[3]{x_0}$ so $F(x)$, $F(x^2)$, and $F(x^3)$ will all convergeif and only if $x \le x_0$. Then we can conclude that $R(x)$ converges for all $x \le x_0$. Incedentally this shows that conditions 1, 2, 3, 4, hold. So we may apply Theorem 1 and expand R(x) as

$$R(x) = R(x_0) + \sum_{n=1}^{\infty} b_k (x_0 - x)^{\frac{k}{2}} \tag{29}$$

4

We wish to find the value of b_1 to generate our asymptotic formula. To do this we will use Eq. 27 to relate Eq. 29 to Eq. 18. More specically, we will differentiate Eq. 27 to relate the the differentiated expansions. When we differentiate Eq. 27 we get:

$$R'(x) = \frac{F'(x)}{x} - \frac{F(x)}{x^2} + \frac{F^2(x)F'(x)}{2x^2} - \frac{F^3(x)}{3x^2} + \frac{F'(x)F(x^2)}{2x^2} + \frac{F(x)F'(x^2)}{x} - \frac{F(x)F(x^2)}{x^3} + F'(x^3) - \frac{2F(x^3)}{3x^3} \quad (30)$$

If we were to differentiate Eq. 29 and Eq. 18 we get

$$R'(x) = -\frac{1}{2}b_1(x_0 - x)^{-1/2} - b_2 - \frac{3}{2}b_3(x_0 - x)^{1/2}... \quad (31)$$

and

$$F'(x) = -\frac{1}{2}a_1(x_0 - x)^{-1/2} - a_2 - \frac{3}{2}a_3(x_0 - x)^{1/2}... \quad (32)$$

Notice that as $x \to x_0$ every term goes to zero except for the first two in the two expansions. Additionally, the first term blows up as $x \to x_0$. So we know that $F'(x)$ and $R'(x)$ are diverge at $x = x_0$. Thus as $x \to x_0$ the terms that have a $F'(x)$ will dominate so we can ignore all the other terms. So we write:

$$\lim_{x \to x_0} R'(x) = \lim_{x \to x_0} F'(x) \left(\frac{1}{x} + \frac{F^2(x))}{2x^2} + \frac{F(x^2)}{2x^2} \right)$$

which can be simplified using Eq. 12 as

$$\lim_{x \to x_0} R'(x) = \lim_{x \to x_0} F'(x) \left(\frac{1}{x} + \frac{F(x) - x^2 - xF(x)}{x^2} \right) \quad (33)$$

While $F'(x)$ and $R'(x)$ diverge the portion in the parenthesis converges. So we may substitute in the first term in the expansions for $F'(x)$ and $R'(x)$ into Eq. 33 to get that

$$\lim_{x \to x_0} \frac{1}{2}b_1(x_0 - x)^{-1/2} = \lim_{x \to x_0} \frac{1}{2}a_1(x_0 - x)^{-1/2} \left(\frac{1}{x_0} + \frac{F(x_0) - x_0^2 - x_0 F(x_0)}{x^2} \right)$$

We can cancel and simplify using Eq. 17 to get

$$b_1 = a_1 \left(\frac{1 - x_0}{x_0^2} \right) \quad (34)$$

and we get the following expression for our Asymptotic expression for rooted trees.

$$R_n \approx \left(\frac{1 - x_0}{x_0^2} \right) a_1 \sqrt{\frac{x_0}{2\pi}} x_0^{-n} n^{-3/2} \approx \left(\frac{1 - x_0}{x_0^2} \right) F_n \quad (35)$$

and we can approximate these constants and get

$$R_n \approx 1.174148(.4026975)^{-n} n^{-3/2} \quad (36)$$

3   General Trees

Finally, we will apply this theorem to general trees to get an Asymptotic formula. The functional equation for general trees is

$$T(x) = R(x) - \frac{F^2(x)}{2x^2} + \frac{F(x^2)}{2x^2} \quad (37)$$

At this point it is obvious that all 6 conditions are met and that $T(x)$ converges if and only if $x \leq x_0$ because $T(x) = T(F(x), F(x^2), F(x^3), x)$. So we may immediately apply our Theorem to get the expansion:

$$T(x) = T(x_0) + \sum_{n=1}^{\infty} c_k (x_0 - x)^{\frac{k}{2}} \quad (38)$$

We will find c_1 in the same way we found b_1. We will relate it to b_1 and a_1. We first differentiate Eq. 37 to get:

$$T'(x) = R'(x) - \frac{F(x)F'(x)}{x^2} + \frac{F^2(x)}{x^3} + \frac{F'(x^2)}{x} - \frac{F(x^2)}{x^3} \quad (39)$$

Then we can put in the expansions and take the limit as $x \to x_0$ to get:

$$\frac{c_1}{2}(x_0 - x)^{-1/2} = \frac{b_1}{2}(x_0 - x)^{-1/2} - \frac{a_1 F(x_0)}{2x_0^2}(x_0 - x)^{-1/2}$$

and then we simplify to get

$$c_1 = b_1 - \frac{a_1(1 - x_0)}{x_0^2}$$

and then if we remember Eq. 34 then clearly

$$c_1 = 0 \tag{40}$$

Then according to Theorem 1 we must use the second approximation using c_3. So our task now becomes finding c_3. We will do this by doing a few clever tricks and then taking the second derivaitve of $T(x)$ to get c_3 perfectly interms of a_1. We start by using Eq.30 to rewrite Eq. 39 as

$$
\begin{aligned}
T'(x) =& \frac{F'(x)}{x} - \frac{F(x)}{x^2} + \frac{F^2(x)F'(x)}{2x^2} - \frac{F^3(x)}{3x^2} + \frac{F'(x)F(x^2)}{2x^2} + \frac{F(x)F'(x^2)}{x} - \frac{F(x)F(x^2)}{x^3} \\
&+ F'(x^3) - \frac{2F(x^3)}{3x^3} - \frac{F(x)F'(x)}{x^2} + \frac{F^2(x)}{x^3} + \frac{F'(x^2)}{x} - \frac{F(x^2)}{x^3}
\end{aligned}
\tag{41}
$$

We can gather all of the terms with an $F'(x)$ in them and write

$$T'(x) = F'(x)\left(\frac{1}{x} + \frac{1}{x^2}\left(\frac{F^2(x)}{2} + \frac{F(x^2)}{2} - F(x)\right)\right) + ...$$

which can be simplifed using Eq. 12 as

$$T'(x) = F'(x)\left(\frac{1 - x - F(x)}{x}\right) + ...$$

and finally we can see that we can simplify Eq. 22 be subtracting $xF'(x)$ from both sides to get

$$F'(x)(1 - x - F(x)) = 2x + F(x) + xF'(x^2) \tag{42}$$

We can plug this in to get an expression for T'(x) with no mention of F'(x).

$$
\begin{aligned}
T'(x) =& 2 + \frac{F(x)}{x} + F'(x^2) - \frac{F(x)}{x^2} - \frac{F^3(x)}{3x^2} + \frac{F(x)F'(x^2)}{x} - \frac{F(x)F(x^2)}{x^3} \\
&+ F'(x^3) - \frac{2F(x^3)}{3x^3} + \frac{F^2(x)}{x^3} + \frac{F'(x^2)}{x} - \frac{F(x^2)}{x^3}
\end{aligned}
\tag{43}
$$

Now we may take the second derivative. Since there are no $F'(x)$ terms we will not get any $F''(x)$ terms in the second derivative. This means that all terms in the second derivative will converge as $x \to x_0$ except those containing $F'(x)$ (where as if there were terms with $F''(x)$ these two would diverge complicating our ananlysis). This means that when we look at what happens as $x \to x_0$ we will only care about the divergent $F'(x)$ terms. So for simplicities sake I will leave out all terms in the second derivative that dont contain an $F'(x)$. We get

$$T''(x) = \frac{F'(x)}{x} - \frac{F'(x)}{x^2} - \frac{F^2(x)F'(x)}{x^3} + \frac{F'(x)F'(x^2)}{x} - \frac{F'(x)F(x^2)}{x^3} + \frac{2F(x)F'(x)}{x^3} \tag{44}$$

which can be rewritten as

$$T''(x) = F'(x)\left(\frac{1}{x}(1 + F'(x^2)) - \frac{1}{x^2} + \frac{2}{x^3}\left(F(x) - \frac{F^2(x)}{2} - \frac{F(x^2)}{2}\right)\right) + ... \tag{45}$$

Then we may take the limit as $x \to x_0$ and substitute in Eq. 12 and Eq. 24 to get

$$\lim_{x \to x_0} T''(x) = \lim_{x \to x_0} F'(x)\left(\frac{1}{x_0^2}\left(\frac{a_1^2}{2} - 2\right) + \frac{2}{x_0^3}(x_0 + x_0 F(x_0))\right)$$

6

and the we may substitute in the expansions for $T''(x)$ and $F'(x)$ to get

$$\frac{3c_3}{4}(x_0 - x)^{-1/2} = \left(\frac{a_1}{2}(x_0 - x)^{-1/2}\right)\left(\frac{\frac{a_1^2}{2} - 2 + 2x_0 + 2 - 2x_0}{x_0^2}\right)$$

which can be simplified to simply

$$c_3 = \frac{a_1^3}{3x_0^2} \tag{46}$$

Then, according to Theorem 1 our asymptotic formula for T_n is

$$T_n \approx \frac{a_1^3}{4\sqrt{\pi x_0}}x_0^{-n}n^{-5/2} \tag{47}$$

Then, if we plug in our values for a_1 and x_0 becomes

$$T_n \approx 1.255109(.4026975)^{-n}n^{-5/2} \tag{48}$$

4 Entropy

We are finally in a position where we may extract som thermodynamic variables. Specifically we can immediately calculate entropy. The it is well known that the Entropy of a system may be calculated as

$$S = k_B \ln \Omega \tag{49}$$

where Ω is the multiplicity. What we have been doing this whole time, enumerating trees, is really just calculating the multiplicity of trees with n nodes. Now that we have the asymptotic formulas rather than the recursive ones, we have the multiplicity in a form that fits very nicely into Eq. 49. So if we plug in Eq. 26, Eq. 36, and Eq 48, we can get expressions for the entropy of Planted, Rooted, and most importantly General Trees respectively. after simplfication we arrive at

$$S_{PT} = k_B\left(.90957n - \frac{3}{2}\ln n - 1.14326\right) \tag{50}$$

$$S_{RT} = k_B\left(.90957n - \frac{3}{2}\ln n + .160543\right) \tag{51}$$

$$S_T = k_B\left(.90957n - \frac{5}{2}\ln n + .227222\right) \tag{52}$$

It is useful to note the accuracy of these approximations. For general trees the error in the approximation for $n = 1000$ is $.180727\%$. Thus the error in the entropy is $+k_B \ln 1.00180727$ which is $.00187k_B$. If we use our approximation to the entropy to calculate the entropy for general trees with 1000 nodes we get

$$S_T = 892.5278k_B$$

and then we can easily calulate the percent error in the entropy to be $.0002095\%$ for $n = 1000$. As we increase the node count this error will obviously decrease so we can conclude that:

If $n \geq 1000$ the percent error in the entropy for general trees is less than $.00021\%$