Lengths of Paths In Ordered Trees

Keith Copenhaver

Abstract. We provide formulas for generating functions of many types of paths in various rooted, plane tree structures. We use these to find closed form and asymptotic expectations, including some surprisingly simple closed forms. 

Mathematics Subject Classification 05A15 · 05A16 · 05C05 · 05C12
1. Introduction

A tree is a connected graph without cycles. A tree is rooted if it has one vertex designated as the root, and it is plane if the order of its neighbors matters. In a rooted tree there is a natural orientation imposed on the vertices, the neighbor of a vertex which is closer to the root is called the parent, and the vertices which are further away are called the children. Extending the metaphor, if a vertex $v$ is contained in the subtree rooted at a vertex $u$, then $v$ is a descendant of $u$, and $u$ is an ancestor of $v$.

Trees have widespread applications. They are used to model networks of many kinds, and they are used as data structures. In some of these applications, leaves (vertices which have no children) have particular significance. In data structures, the leaves will always be the last thing called in a search algorithm, so one might like the leaves to be relatively close to the root. If we represent a computer network as a rooted tree, then the leaves are the nodes with the least amount of access, and presumably, the least security; in this application, we might want most vertices to be far from leaves. Height is the length of the longest path from the root; this path necessarily ends at a leaf. Height has been a longstanding topic of study. There have been a number of papers recently about protection and rank, starting with the seminal paper by Cheon and Shapiro \[1\]. A vertex is $k$-protected if its closest leaf descendant is at least $k$ steps away, and it has rank $k$ if its closest leaf descendant is exactly $k$ steps away. Height and rank provide two halves of the same coin: on some level, they both address the question “how far away are the leaves from the root?” It is a decades old result from de Bruijn, Knuth, and Rice that the furthest leaf in a uniformly selected general tree of arbitrarily large size is expected to be arbitrarily far away \[3\]. This was later proven for binary trees by Flajolet and Odlyzko as well \[5\]. It is a recent result that the closest leaf in general trees is expected to be less than two edges away \[2\]. This leaves the question of what exactly happens in the middle? We will answer that question precisely with both explicit formulas and asymptotic estimations.

This also leads to generalization: what if one were to pick two vertices arbitrarily? What if we pick a vertex and any of its descendants? We will call paths between a vertex and one of its descendants “downward paths.”

We will study two types of unlabeled, rooted plane trees here. General trees are trees in which every vertex may have any number of children. Binary trees are trees in which every vertex may have zero, one, or two children, where single children may be left or right children. There are some interesting relationships in these two families, for example:

- The expected length of a path from the root to a leaf is very close to the average of the expected distances to the closest and furthest leaves.
- Even downward paths, the paths which have the shortest expected length, are expected to be arbitrarily long in arbitrarily large trees.
- In these families of trees, insisting that paths end in leaves does not result in large difference in the expected length of downward paths; asymptotically the expectations differ by only a constant.
• There is a great deal of overlap across the generating functions for both number of paths and number of edges in paths within each family. This situation is less pronounced in other families of trees, where the generating function for the family is less pleasant.
• The expected length of a path from a root to a leaf is asymptotically equal to the expected length of an arbitrary path in trees on \( n \) vertices, and, in the case of general trees, these expectations are exactly equal for all \( n \).

2. Grafting

In horticulture, grafting is a process where a branch of one tree is removed and made to grow as a branch of a different tree. We will use a lemma which does something similar with graph theoretic trees. We will call a family of trees \( \mathcal{T} \) \textit{graftable} if the set of subtrees which may be rooted at any vertex is equal to the set of trees.

For example, most rooted, plane, unlabeled tree structures are graftable, since any subtree is a valid tree in its own right. This includes any tree where, given some subset of the positive integers \( \Omega \) with \( 0 \in \Omega \), any tree can be decomposed as a root and a sequence of \( k \) trees, where \( k \in \Omega \), or in the symbolic method, \( T^\Omega = z \times \text{SEQ}_\Omega (T^\Omega) \). Such trees are referred to as \( \Omega \)-restricted trees in Chapter I.5 of [6]. However, consider trees where the root may only have an odd number of children but all other vertices may have any number of children. These trees are not graftable, because any subtree with an even number of children is not a valid tree.

**Lemma 2.1 (Grafting Lemma).** Let \( \mathcal{T} \) be a graftable class of trees. Then there is a bijection between the set of vertices in trees of size \( n \) in \( \mathcal{T} \) and the set of ordered pairs of leaves in trees of size \( n - k + 1 \) and trees of size \( k \), with \( 1 \leq k \leq n \). Further, this bijection extends to allow marking vertices or edges and restricts to subsets of subtrees and trees.

**Proof.** Fix a vertex and remove the subtree rooted at that point, replacing it with a leaf. This gives a leaf and a tree. Conversely, fix a leaf and a tree, remove the leaf, and replace it by placing the tree at that vertex as a subtree. Note that if there were any marked edges or vertices in the tree, they will now correspond to marked vertices or edges in the designated subtree. \( \square \)

On a basic level, this implies that

(2.1) \[ V_\mathcal{T} = \frac{L_\mathcal{T}(x)}{x} T_\mathcal{T}(x), \]

and, since

\[ V_\mathcal{T}(x) = \sum_{n=1}^{\infty} n ([x^n] T_\mathcal{T}(x)) x^n = x T'_\mathcal{T}(x), \]

rearranging yields

\[ L_\mathcal{T}(x) = \frac{x^2 T'_\mathcal{T}(x)}{T_\mathcal{T}(x)}. \]
This gives a method for computing the generating function for leaves in any case where the generating function for trees is known, without requiring bivariate generating functions or insights into the functional equation satisfied by the trees. However, if it is known that the generating function satisfies a functional equation of the form $T_T = x\Phi_T(T_T)$, this relation implies that $xT'_T(x) = T_T(1 - x\Phi_T(T_T))^{-1}$, (this equation was noted in [4]) and comparing these equations yields

$$L_T(x) = \frac{x}{1 - x\Phi_T(T_T(x))}.$$  

This bijection is powerful because it preserves many qualities of the tree (or root) being grafted. It is generally simpler to construct a generating function for trees with roots of a given type, and this immediately generalizes those results to arbitrary vertices. The core idea of this lemma was used in [1].

**Theorem 2.2.** Let $T$ be a graftable class of trees with generating functions $T_T(x), V_T(x),$ and $L_T(x)$ for the number of trees, vertices, and leaves on $n$ vertices, respectively, where empty trees are not allowed. Then, for all trees on $n$ vertices:

1. The number of downward paths from the root has generating function $V_T(x) - T_T(x)$.
2. The number of all downward paths has generating function $\frac{L_T(x)}{x} \cdot (V_T(x) - T_T(x))$.
3. The number of edges in all downward paths originating at the root also has generating function $\frac{L_T(x)}{x} \cdot (V_T(x) - T_T(x))$.
4. The number of edges in all downward paths has generating function $\left(\frac{L_T(x)}{x}\right)^2 \cdot (V_T(x) - T_T(x))$.
5. For (1)–(4), if we insist that the path terminate at a leaf, replace $(V_T(x) - T_T(x))$ with $(L_T(x) - [x]T_T(x))$ (here $[x]T_T(x)$ indicates the coefficient of $x$ in the generating function $T_T(x)$).

Each of these parts is provable using bivariate generating functions and algebraic manipulation, but the grafting lemma provides a straightforward univariate proof which is contained in the appendix. The number of edges in all downward paths from the root in a tree is called the path length of the tree, and it is discussed extensively in [6].

### 2.1. Downward Paths in General Trees.

Let $G$ be the class of general. Then the generating functions for trees, vertices, and leaves, respectively, are

$$T_G(x) = \frac{1 - \sqrt{1 - 4x}}{2}, \quad V_G(x) = \frac{x}{\sqrt{1 - 4x}}, \quad \text{and} \quad L_G(x) = \frac{2x + x^2}{2\sqrt{1 - 4x}}.$$  

By Theorem 2.2, the generating function for the number of paths from the root to a leaf in trees on $n$ vertices is

$$\left(\frac{1}{2\sqrt{1 - 4x}} + \frac{1}{2}\right) \left(\frac{x}{2\sqrt{1 - 4x}} + \frac{x}{2} - x\right) = \frac{x^2}{1 - 4x}.$$
Undergraduate calculus shows that the number of edges in such paths is $4^{n-2}$. Since the number of such paths is clearly equal to the number of leaves, $\frac{1}{2}(2^{n-2})$, for $n \geq 2$ (each leaf has a unique path to the root), it follows that the expected length of such a path is

$$\frac{4^{n-2}}{\binom{2(2n-2)}{n-1}} = \frac{4^{n-2}2^{n-3}2(2n-3)!}{2(2n-2)!} = \frac{(2n-2)!}{2(2n-3)!} \approx \frac{\sqrt{\pi n}}{2} + O\left(\frac{1}{\sqrt{n}}\right).$$

A better asymptotic approximation for this number is given in the appendix. Since the expected distance to the closest leaf is approximately $\frac{1}{2} + O\left(\frac{1}{n}\right)$, and the furthest leaf is $\sqrt{\pi n} - \frac{1}{2} + O\left(\frac{1}{n}\right)$, the difference between the expected distance from the root to a leaf and the average of the distances to the closest and furthest leaves is only about $0.561486$ in an arbitrarily large tree.

By Theorem 2.2, the generating function for the number of edges in downward paths in general trees is

$$\left(\frac{1}{2\sqrt{1-4x}} + \frac{1}{2}\right)^2 \left(\frac{x}{\sqrt{1-4x}} - \frac{1 - \sqrt{1-4x}}{2}\right) = \frac{x^2}{(1-4x)^{3/2}},$$

and the generating function for the number of such paths is

$$\left(\frac{1}{2} + \frac{1}{2\sqrt{1-4x}}\right) \left(\frac{x}{\sqrt{1-4x}} - \frac{1 - \sqrt{1-4x}}{2}\right) = \frac{x}{2(1-4x)} - \frac{x}{2\sqrt{1-4x}}.$$

Both of these generating functions are straightforward enough to extract exact coefficients:

$$[x^n] \frac{x^2}{(1-4x)^{3/2}} = \frac{(2n-3)!}{(n-2)!^2}, [x^n] \left(\frac{x}{2(1-4x)} - \frac{x}{2\sqrt{1-4x}}\right) = \frac{1}{2} \left(4^{n-1} - \binom{2n-2}{n-1}\right).$$

The expected length of a downward path in a general tree on $n$ vertices is the ratio of the two,

$$\frac{\frac{(2n-3)!}{(n-2)!^2}}{\frac{1}{2} \left(4^{n-1} - \binom{2n-2}{n-1}\right)} = \frac{n - 1}{\frac{(2n-2)!}{(2n-3)!} - 1}.$$

A careful reader may note that the number of edges in downward paths is equal to $\binom{n}{2} c_{n-1}$, where $c_{n-1} = \frac{1}{n} \binom{2n-2}{n-1}$, the $(n-1)$st Catalan number, which is the number of trees on $n$ vertices. There is also the slightly more obvious fact that the generating function is one half times the second derivative of the generating function for trees. This is also the number of all paths, since one can choose any two endpoints and uniquely determine a path. These generating function arguments prove that their number is the same, but the fact invites a bijective proof.

**2.2. Arbitrary Paths in General Trees.** For this section, it will be useful to repeatedly refer to the edge between the root and its leftmost child. Thus, we will refer to this edge as the “key edge.”

**Proposition 2.3.** In general trees on $n$ vertices, the number of paths is equal to the number of edges in downward paths.
Proof. We construct a bijection between the set of marked edges in marked downward paths to marked paths. Note that this bijection will not, in general, map into the same tree.

Take a path. If the path is downward send it to the same path and mark the top edge. Else, mark the edge in the path on the right of the vertex closest to the root, and remove this edge and all of its siblings to its right. This gives a subtree where there is path through the key edge, and the key edge is marked. Graft this subtree onto the end of the left-hand path, to the right of the existing children; this gives a downward path with a single marked edge.

\[\square\]

Figure 1. An example of the bijection used in Proposition 2.3.

Proposition 2.4. In all general trees on \(n\) vertices, the number of edges in all paths which include the key edge is \(\binom{n}{2}c_{n-1}\).

Proof. We establish a bijection between all paths in trees on \(n\) vertices to edges in paths which pass through the key edge. We will group all paths into one of four cases. There is overlap between cases 1 and 2, but the mappings agree.

Case 1 Paths which contain the key edge: for any such path, map it to the same path in the same tree, and, within that path, mark the key edge.

Case 2 Downward paths: for any such path, map it to the unique path which passes from the key edge to the bottom edge of this path. This path passes through the root if the downward path is not contained in the subtree beneath the key edge. Next, mark the edge which was the edge of the original path which was closest to the root.

Case 3 Non-downward paths which are contained in the subtree below the key edge: consider the vertex closest to the root; remove the subtree whose key edge is the right-hand edge of the path. This leaves two trees with vertical paths, the second with the path through its key edge (this is necessary for the reversal of the algorithm). Mark the top edge of the path in the first tree. Graft the second tree to the root of the first tree, to the right of all existing edges. Finally, connect the two paths.
Case 4 Non-downward paths which are not contained in the subtree below the key edge: similarly to case 3, take the vertex closest to the root and remove the subtree whose key edge is the first edge of the right-hand path. Mark the top edge of the path that remains, then graft the second tree at the bottom of the key edge, to the right of existing vertices, then connect.
This theorem stands apart from Theorem 2.2, despite the fact that it seems to correspond to the relationship between 2.2 (3) and 2.2 (4). Both theorems establish that paths anywhere in the tree in some way correspond to edges in paths through the root, but Theorem 2.2 applies to any class of graftable trees, and relies on straightforward bijections. This theorem only applies to general trees and uses heavily the fact that general trees admit any number of children.

Corollary 2.5. The number of edges in all paths, in trees on $n$ vertices is $(n − 1)4^{n−2}$. Further, the expected length of a uniformly selected path in a tree of size $n$ is equal to the expected length of a uniformly selected path from the root to a leaf in trees of size $n$.

Proof. We may decompose any path with a marked edge into a tree with a marked vertex and a tree with a path through the key edge with a marked edge. This is done as follows: take the vertex which is closest to the root contained in the path and remove the subtree starting at the left edge of the path.
Let $E_G(x)$ be the generating function for the number of edges in all paths on $n$ vertices. Then, by the product formula for generating functions,

$$E_G(x) = \frac{1}{x}V_G(x) \cdot \frac{x^2}{(1-4x)^{3/2}} = \frac{x^2}{(1-4x)^{2}},$$

and the result follows by the binomial theorem.

Thus the expected length of a uniformly randomly selected path in a uniformly randomly selected ordered tree on $n$ vertices is

$$\mathbb{E} [\text{length}] = \frac{(n-1)4^{n-2}}{\binom{n}{2} \binom{2n-2}{n-1}} = \frac{(n-1)4^{n-2}}{\binom{n}{2} \binom{2n-3}{n-2}} = \frac{2n-2)!}{2(2n-3)!}.$$ 

Thus, for all $n \geq 2$, the expected length of a uniformly selected path is equal to the expected length of a uniformly selected path from the root to a leaf (corresponding with (2.2)).

The sum of all distances in a graph is also called the Wiener Index of a graph. The Wiener Index for general trees was previously derived by Entringer, Meir, Moon, and Székely using different methods, including a nice bijective proof [4].

**Proposition 2.6.** The number of paths between leaves in all general trees on $n$ vertices is

$$\binom{n}{2} \frac{(n-1)4^{n-2}}{\binom{2n-2}{n-1}} = \frac{(n-1)4^{n-2}}{\binom{n}{2} \binom{2n-3}{n-2}} = \frac{2n-2)!}{2(2n-3)!}.$$ 

Thus, for all $n \geq 2$, the expected length of a uniformly selected path is equal to the expected length of a uniformly selected path from the root to a leaf (corresponding with (2.2)).

The sum of all distances in a graph is also called the Wiener Index of a graph. The Wiener Index for general trees was previously derived by Entringer, Meir, Moon, and Székely using different methods, including a nice bijective proof [4].

**Proposition 2.6.** The number of paths between leaves in all general trees on $n$ vertices is

$$\binom{n}{2} \frac{(n-1)4^{n-2}}{\binom{2n-2}{n-1}} = \frac{(n-1)4^{n-2}}{\binom{n}{2} \binom{2n-3}{n-2}} = \frac{2n-2)!}{2(2n-3)!}.$$ 

Thus, for all $n \geq 2$, the expected length of a uniformly selected path is equal to the expected length of a uniformly selected path from the root to a leaf (corresponding with (2.2)).

The sum of all distances in a graph is also called the Wiener Index of a graph. The Wiener Index for general trees was previously derived by Entringer, Meir, Moon, and Székely using different methods, including a nice bijective proof [4].

**Proposition 2.6.** The number of paths between leaves in all general trees on $n$ vertices is

$$\binom{n}{2} \frac{(n-1)4^{n-2}}{\binom{2n-2}{n-1}} = \frac{(n-1)4^{n-2}}{\binom{n}{2} \binom{2n-3}{n-2}} = \frac{2n-2)!}{2(2n-3)!}.$$ 

Thus, for all $n \geq 2$, the expected length of a uniformly selected path is equal to the expected length of a uniformly selected path from the root to a leaf (corresponding with (2.2)).

The sum of all distances in a graph is also called the Wiener Index of a graph. The Wiener Index for general trees was previously derived by Entringer, Meir, Moon, and Székely using different methods, including a nice bijective proof [4].

**Proposition 2.6.** The number of paths between leaves in all general trees on $n$ vertices is

$$\binom{n}{2} \frac{(n-1)4^{n-2}}{\binom{2n-2}{n-1}} = \frac{(n-1)4^{n-2}}{\binom{n}{2} \binom{2n-3}{n-2}} = \frac{2n-2)!}{2(2n-3)!}.$$ 

Thus, for all $n \geq 2$, the expected length of a uniformly selected path is equal to the expected length of a uniformly selected path from the root to a leaf (corresponding with (2.2)).

The sum of all distances in a graph is also called the Wiener Index of a graph. The Wiener Index for general trees was previously derived by Entringer, Meir, Moon, and Székely using different methods, including a nice bijective proof [4].
10 KEITH COPENHAVER

terminates at a leaf, and adding a single vertex to the root, on the left. Thus, by Theorem 2.2, the generating function for such cases is

\[
x \cdot \frac{L_G(x)}{x} (L_G(x) - x) = \frac{x^3}{1 - 4x}.
\]

By the same argument as used in the Corollary 2.5, the generating function for edges in paths from leaves is

\[
\frac{V_G(x)}{x} \left( \frac{x^3}{(1 - 4x)^{3/2}} + \frac{x^3}{1 - 4x} \right) = \frac{x^3}{(1 - 4x)^2} + \frac{x^3}{(1 - 4x)^{3/2}},
\]

and the result follows by extracting coefficients. □

This implies that the number of edges in paths between leaves in trees on \(n\) vertices is equal to the number of edges in all paths plus the number of paths, or, equivalently, the number of all marked vertices in all marked paths in trees on \(n - 1\) vertices. This result also invites a bijective proof, which is not known to the author.

**Corollary 2.7.** The expected length of a uniformly selected path between leaves in a uniformly selected general tree on \(n\) vertices is

\[
(2.5) \quad \frac{(2n - 4)!}{2(2n - 5)!} + 1 \approx \frac{\sqrt{\pi n}}{2} + 1 - \frac{7\sqrt{\pi}}{16\sqrt{n}} + O\left(\frac{1}{n^{3/2}}\right).
\]

The exact expression is attained by taking ratios, and the asymptotic expression is attained by using singularity analysis on the generating functions and once again taking ratios.

### 2.3. Arbitrary Paths in Binary Trees.

Let \(B\) be the class of binary trees. First, let \(p_{n,k}\) be the number of paths in trees on \(n\) vertices which originate at the root and end at a a leaf, and let

\[
P_B(x, y) = \sum_{n,k \geq 1} p_{n,k} x^n y^k.
\]

If we take such a tree, either it is a single vertex, or we can remove the root, leaving a marked edge and a similar tree in one of two ways (since left or right matters), or we may have a similar tree and a regular binary tree in one of two ways. Thus

\[
P_B(x, y) = x + 2xyP_B(x, y) + 2xyP_B(x, y)T_B(x),
\]

with \(T_B(x) = \frac{1 - 2x - \sqrt{1 - 4x}}{2x}\). Solving for \(P(x, y)\),

\[
P(x, y) = \frac{x}{1 - y + y\sqrt{1 - 4x}}.
\]

With this in hand, we can find the bivariate generating function for all paths in a tree using the grafting lemma. First, we graft a tree onto the end of our marked path so that the generating function for paths that start at the root and end anywhere is \(\frac{P(x,y)}{x} \cdot T_B(x)\). The generating function for a path that passes through the root is \(xy^2 \left(\frac{P(x,y)}{x} \cdot T_B(x)\right)^2\), since, if we remove the root, the tree must have two marked edges leading to two children, both of which are a tree with a marked path from the root. The sum of these two gives all paths whose top vertex is the root. We then apply the grafting lemma to give paths whose top vertex is
anywhere in the tree. Let \( p_{n,k}^* \) be the number of all paths in binary trees on \( n \) vertices with \( k \) edges, and let

\[
P_B^*(x,y) = \sum_{n,k \geq 1} p_{n,k}^* x^n y^k.
\]

Then

\[
P^*(x,y) = P(x,y) \cdot T_B(x) + xy^2 \left( \frac{P(x,y)}{x} \cdot T_B(x) \right)^2.
\]

Then the generating function for the number of edges of trees on size \( n \) is

\[
\frac{\partial}{\partial y} P^*(x,y) \bigg|_{y=1} = \frac{1 - \sqrt{1 - 4x - 2x(1 - 4x)}}{(1 - 4x)^2}.
\]

The generating function for the number of such paths is simply \( \frac{x^2}{2} \frac{d^2}{dx^2} T_B(x) \), so, applying singularity analysis in both cases and taking ratios, we have that the expected length of an a uniformly selected path in a binary tree on \( n \) vertices is

\[
\sqrt{\pi n} - 4 + O \left( \frac{1}{n^{1/2}} \right).
\]

To find the generating functions for arbitrary paths which end in leaves, we carry out the same process, but omit the multiplication by \( T_B(x) \) at the end of the paths, which gives the generating functions and results supplied in the appendix. It is straightforward to show that the generating function for the number of paths between leaves is \( x(L_B(x))^2 \).

3. Appendix

3.1. Grafting Lemma.

**Proof of Theorem 2.2.**

1. There is a unique path between any two vertices in a tree of any kind; therefore we need only pick a vertex which is not the root, so the number of these is the number of vertices minus the number of roots; since there is one root in any tree the proposition follows.

2. If we take a tree with a marked path downward path from the root, we may graft it to any leaf and have a tree with a marked downward path at an arbitrary position.

3. Take an arbitrary marked downward path. Color its top edge, then connect the path to the root (if it does not already originate at the root). This gives a bijection between arbitrary downward paths and edges in downward paths which originate at the root.

4. Graft a tree with a marked edge in a marked downward path from the root to an arbitrary leaf to give marked edges in marked downward paths anywhere.

5. The arguments are identical for paths which end in leaves, except with arbitrary vertices replaced by leaves; thus we use \( L(x) \) instead of \( V(x) \). In this case, to avoid choosing a leaf that is the root, we remove only the singleton case, since otherwise, the root cannot be a leaf; thus we replace \( T(x) \) with the number of trees on a single vertex, \( [x]T_T(x) \).
Each piece of the theorem can also be treated more independently, and this does give more visual proofs. For example, for part (4), take an ordered triple of two trees with marked leaves and one tree with a marked vertex to give a downward path with a marked edge as in Figure 1.

![Figure 7](image_url)

**Figure 7.** An example corresponding to Theorem 2.2 (4), where the path with the marked edge may be reconstructed from the ordered triple of two trees with marked leaves and one with a marked vertex. The first marked leaf is the start of the path, the second marked leaf is the top of the marked edge, and the marked vertex is the end of the path.

### 3.2. Calculated results.

Asymptotic approximations on this table were calculated using singularity analysis (see [5], Ch. VI). One example of such calculations is included here, and almost all of the analysis for the others is similar, though in many cases, more tedious.

The generating functions for the number of edges in downward paths in general trees is 
\[
x^2 \binom{1}{(1 - 4x)^{3/2}},
\]
and the generating function for the number of such paths is
\[
x \frac{x^2}{2(1 - 4x)} - \frac{2 \sqrt{1 - 4x}}{2\sqrt{1 - 4x}}.
\]
Both of these functions are analytic on a $\Delta$-domain about $x = 1/4$, so we may apply singularity analysis. We express both functions as power series about $x = 1/4$, so we make the substitution $x = \frac{1}{4}(1 - (1 - 4x))$, yielding

\[
[x^n] \frac{x^2}{(1 - 4x)^{3/2}} = [x^n] \frac{1}{16} \left( \frac{1}{2} \left( 1 + \frac{1}{2} \right) - \frac{2}{\sqrt{1 - 4x}} + \sqrt{1 - 4x} \right)
\]

\[
= 4^n \left( \frac{\sqrt{n}}{\sqrt{\pi}} + \frac{1}{4n} - \frac{1}{64n^2} + O \left( \frac{1}{n^3} \right) \right) - \frac{2}{\sqrt{\pi}n^3} \left( 1 - \frac{1}{8n} + O \left( \frac{1}{n^2} \right) \right) - \frac{1}{\sqrt{\pi}n^3} \left( \frac{1}{2} + O \left( \frac{1}{n} \right) \right)
\]

\[
= 4^n \left( \frac{\sqrt{n}}{8\sqrt{\pi}} - \frac{5}{64\sqrt{\pi}n} - \frac{23}{1024\sqrt{\pi}n^3} \right) + O \left( \frac{4^n}{n^{5/2}} \right).
\]
\[ [x^n] \frac{x}{2(1-4x)} - \frac{x}{2\sqrt{1-4x}} = [x^n] \frac{1}{8} \left( \frac{1}{1-4x} - \frac{1}{\sqrt{1-4x}} - 1 + \sqrt{1-4x} \right) \]
\[ = \frac{4^n}{8} \left( 1 - \frac{1}{\sqrt{\pi n}} \left( 1 - \frac{1}{8n} + O \left( \frac{1}{n^2} \right) \right) \right) \]
\[ - \frac{1}{\sqrt{\pi n^3}} \left( \frac{1}{2} + O \left( \frac{1}{n} \right) \right) \]
\[ = 4^n \left( \frac{1}{8} - \frac{1}{8\sqrt{\pi n}} - \frac{3}{64\sqrt{\pi n^3}} \right) + O \left( \frac{4^n}{n^{8/2}} \right). \]

Taking ratios, we have that the expected length of a uniformly selected downward path among all trees on \( n \) vertices is
\[ \frac{\sqrt{n}}{\sqrt{\pi}} + \frac{1}{\pi} \frac{8 - 5\pi}{8\pi^{3/2} \sqrt{n}} + \frac{4 - \pi}{4\pi^2 n} + \frac{128 + 16\pi - 23\pi^2}{128\pi^{5/2} n^{3/2}} \]
\[ + \frac{32 + 16\pi - 7\pi^2}{32\pi^3 n^2} + O \left( \frac{1}{n^{5/2}} \right). \]

All formulas on the following tables should be taken for \( n \geq 2 \), or in the cases of paths between leaves, for \( n \geq 3 \), since there are no paths on trees of lesser size. Thus, some of the formulas do not match the coefficients for the generating functions for smaller values of \( n \).
| General Trees | Number | Generating Function | Coefficient | Asymptotic Coefficient |
|---------------|--------|---------------------|-------------|-----------------------|
| Downward Paths from The Root |        | $\frac{x}{\sqrt{1-4x}} \quad \frac{1 - \sqrt{1-4x}}{2}$ | $\binom{2n-2}{n}$ | $4^n \left( \frac{1}{4\sqrt{\pi n}} - \frac{5}{32\sqrt{\pi n^3}} - \frac{23}{512\sqrt{\pi n^5}} \right) + O \left( \frac{4^n}{n^{7/2}} \right)$ |
| Edges |        | $\frac{x}{2(1-4x)} \quad \frac{x}{2\sqrt{1-4x}}$ | $\frac{1}{2} \left( 4^{n-1} - \frac{2n-2}{n-1} \right)$ | $4^n \left( \frac{1}{8\sqrt{\pi n}} - \frac{1}{64\sqrt{\pi n^3}} \right) + O \left( \frac{4^n}{n^{7/2}} \right)$ |
| Expected Length |        | | | $n \left( \frac{2n-4}{(2n-3)!!} - \frac{1}{2(n-1)} \right)$ | $\frac{\sqrt{\pi n}}{2} - \frac{1}{2} + \frac{5\sqrt{\pi}}{16\sqrt{n}} - \frac{1}{2n} + \frac{73\sqrt{\pi}}{256\sqrt{n^3}} - \frac{1}{2n^2} + O \left( \frac{1}{n^{5/2}} \right)$ |
| Downward Paths | | $\frac{x}{(1-4x)^{3/2}}$ | $\binom{2n-3}{(n-2)!!}$ | $4^n \left( \frac{\sqrt{\pi}}{5\sqrt{n}} - \frac{5}{64\sqrt{\pi n^3}} \right) + O \left( \frac{4^n}{n^{7/2}} \right)$ |
| Edges | | $\frac{x^2}{1 - 4x}$ | $4^{n-2}$ | $\frac{\sqrt{\pi n}}{2} - \frac{3\sqrt{\pi}}{16\sqrt{n}} - \frac{7\sqrt{\pi}}{256\sqrt{n^3}} + O \left( \frac{1}{n^{5/2}} \right)$ |
| Expected Length | | | | | $\frac{\sqrt{\pi n}}{2} - \frac{3\sqrt{\pi}}{16\sqrt{n}} - \frac{7\sqrt{\pi}}{256\sqrt{n^3}} + O \left( \frac{1}{n^{5/2}} \right)$ |
| Path Type                      | Number Generating Function | Coefficient | Edge Generating Function | Coefficient | Asymptotic Coefficient |
|-------------------------------|-----------------------------|-------------|--------------------------|-------------|------------------------|
| **General Trees, cont.**      |                             |             |                          |             |                        |
| Downward Paths Which End in Leaves | $x^2$                     | $4^{n-2}$   | $\frac{x^2}{2(1 - 4x)^{3/2}} + \frac{x^2}{2(1 - 4x)}$ | $\frac{(2n - 3)!}{2((n - 2)!)^2} + \frac{4^{n-2}}{2}$ | $4^n \left( \frac{\sqrt{n}}{16\sqrt{\pi}} + \frac{1}{32} - \frac{5}{128\sqrt{\pi n}} - \frac{23}{2048\sqrt{\pi n^3}} \right) + O \left( \frac{4^n}{n^{5/2}} \right)$ |
| Expected Length               |                             |             |                          |             |                        |
|                               | $\frac{x^2}{(1 - 4x)^{3/2}}$ |             |                          |             |                        |
| Arbitrary Paths               |                             |             |                          |             |                        |
|                               | $x^2$                     |             | $\frac{(2n - 3)!}{((n - 2)!)^2}$ |             | $4^n \left( \frac{\sqrt{n}}{8\sqrt{\pi}} - \frac{5}{64\sqrt{\pi n}} - \frac{23}{1024\sqrt{\pi n^3}} \right) + O \left( \frac{4^n}{n^{5/2}} \right)$ |
|                               | $(n - 1)4^{n-2}$           |             |                          |             |                        |
|                                 | $\frac{x^3}{(1 - 4x)^{5/2}}$ |             |                          |             |                        |
| Path Between Leaves           | $x^3$                     |             |                          |             |                        |
|                               | $\frac{x^3}{(1 - 4x)^2} + \frac{x^3}{(1 - 4x)^3/2}$ |             |                          |             |                        |
|                                 | $\frac{(n - 2)4^{n-3} + (2n - 5)!}{((n - 3)!)^2}$ |             |                          |             |                        |
|                                 | $4^n \left( \frac{n}{64} + \frac{\sqrt{n}}{32\sqrt{\pi}} - \frac{1}{32} - \frac{9}{256\sqrt{\pi n}} \right) + O \left( \frac{4^n}{n^{3/2}} \right)$ |
| Expected Length               | $\frac{(2n - 4)!}{2(2n - 5)!} + 1$ |             |                          |             |                        |
|                                 | $\frac{\sqrt{n}}{2} + \frac{7\sqrt{\pi}}{16\sqrt{n}} - \frac{47\sqrt{\pi}}{256n^{3/2}} + O \left( \frac{1}{n^{5/2}} \right)$ |
| Downward Paths from The Root Which End in Leaves | Number | Expected Length |
|-------------------------------------------------|--------|-----------------|
| Generating Function                              | $1 - \frac{1}{2} \sqrt{\frac{1}{4} - 4x} + \frac{1}{x\sqrt{1 - 4x}}$ | $\sqrt{\frac{\pi n}{3}} - 3 + \frac{17\sqrt{\frac{\pi}{4}}}{8\sqrt{\pi n^3}} - \frac{289\sqrt{\frac{\pi}{4}}}{128\sqrt{\pi n^5}} + O\left(\frac{4^n}{n^{5/2}}\right)$ |
| Coefficient                                      | $2n + 2$ | $\frac{n}{n + 1}$ |
| Asymptotic Coefficient                           | $4^n \left(\frac{1}{\sqrt{\pi n}} - \frac{17}{8\sqrt{\pi n^3}} + \frac{289}{128\sqrt{\pi n^5}}\right) + O\left(\frac{4^n}{n^{5/2}}\right)$ | $\sqrt{\frac{\pi n}{3}} - 3 + \frac{17\sqrt{\frac{\pi}{4}}}{8\sqrt{\pi n^3}} - \frac{289\sqrt{\frac{\pi}{4}}}{128\sqrt{\pi n^5}} + O\left(\frac{4^n}{n^{5/2}}\right)$ |
| Edges                                            | $1 - \frac{1}{x(1 - 4x)} - \frac{1}{x\sqrt{1 - 4x}} - \frac{3}{1 - 4x} + \frac{1}{x\sqrt{1 - 4x}}$ | $\sqrt{\frac{\pi n}{3}} - 3 + \frac{17\sqrt{\frac{\pi}{4}}}{8\sqrt{\pi n^3}} - \frac{289\sqrt{\frac{\pi}{4}}}{128\sqrt{\pi n^5}} + O\left(\frac{4^n}{n^{5/2}}\right)$ |
| Coefficient                                      | $2n + 2$ | $\frac{n}{n + 1}$ |
| Asymptotic Coefficient                           | $4^n \left(\frac{1}{\sqrt{\pi n}} - \frac{17}{8\sqrt{\pi n^3}} + \frac{289}{128\sqrt{\pi n^5}}\right) + O\left(\frac{4^n}{n^{5/2}}\right)$ | $\sqrt{\frac{\pi n}{3}} - 3 + \frac{17\sqrt{\frac{\pi}{4}}}{8\sqrt{\pi n^3}} - \frac{289\sqrt{\frac{\pi}{4}}}{128\sqrt{\pi n^5}} + O\left(\frac{4^n}{n^{5/2}}\right)$ |
| Expected Length                                  | $\sqrt{\frac{\pi n}{3}} - 3 + \frac{17\sqrt{\frac{\pi}{4}}}{8\sqrt{\pi n^3}} - \frac{289\sqrt{\frac{\pi}{4}}}{128\sqrt{\pi n^5}} + O\left(\frac{4^n}{n^{5/2}}\right)$ | $\sqrt{\frac{\pi n}{3}} - 3 + \frac{17\sqrt{\frac{\pi}{4}}}{8\sqrt{\pi n^3}} - \frac{289\sqrt{\frac{\pi}{4}}}{128\sqrt{\pi n^5}} + O\left(\frac{4^n}{n^{5/2}}\right)$ |

| Binary Trees | Generating Function |
|--------------|---------------------|
| $x$          | $x\sqrt{1 - 4x}$    |
| Coefficient  | $2n - 2$            |
| Asymptotic Coefficient | $4^n \left(\frac{1}{\sqrt{\pi n}} + \frac{3}{32\sqrt{\pi n^3}} + \frac{25}{512\sqrt{\pi n^5}}\right) + O\left(\frac{4^n}{n^{1/2}}\right)$ |

| Downward Paths from The Root                     |
|-------------------------------------------------|
| Generating Function                              |
| $x$                                              |
| Coefficient                                      |
| $\frac{2n - 2}{n - 1}$                          |
| Asymptotic Coefficient                           |
| $4^n \left(\frac{1}{\sqrt{\pi n}} + \frac{3}{32\sqrt{\pi n^3}} + \frac{25}{512\sqrt{\pi n^5}}\right) + O\left(\frac{4^n}{n^{1/2}}\right)$ |
### Binary Trees, cont.

| Number | Generating Function | Coefficient | Asymptotic Coefficient |
|--------|---------------------|-------------|------------------------|
| Downward Paths Which End in Leaves | $x/(1 - 4x)$ | $4^n - 3$ | $4^n \left( \frac{\sqrt{n}}{2\sqrt{\pi n}} - \frac{1}{16\sqrt{\pi n}} - \frac{1}{256\sqrt{\pi n^3}} + O\left(\frac{4^n}{n^{5/2}}\right) \right)$ |
| Edges | $x(1 - 4x)^{3/2}$ | $2(2n - 5)!/(n - 3)!^2$ | $4^n \left( \frac{\sqrt{n}}{16\sqrt{\pi n}} - \frac{9}{128\sqrt{\pi n}} - \frac{79}{2048\sqrt{\pi n^3}} + O\left(\frac{4^n}{n^{5/2}}\right) \right)$ |
| Arbitrary Paths | $x^3/(1 - 4x)$ | $4^n - 3$ | $4^n \left( \frac{\sqrt{n}}{2\sqrt{\pi n}} - \frac{9}{32\sqrt{\pi n^3}} - \frac{79}{2048\sqrt{\pi n^3}} + O\left(\frac{1}{n^{5/2}}\right) \right)$ |
| Edges | $x(1 - 4x)^{3/2}$ | $2(2n - 5)!/(n - 6)!^2$ | $4^n \left( \frac{\sqrt{n}}{2\sqrt{\pi n}} - \frac{9}{32\sqrt{\pi n^3}} - \frac{79}{2048\sqrt{\pi n^3}} + O\left(\frac{1}{n^{5/2}}\right) \right)$ |

### Lengths of Paths in Ordered Trees

- **Expected Length**: $\frac{\sqrt{n}}{\sqrt{\pi n}} - \frac{9}{2\sqrt{\pi n^3}} - \frac{79}{32\sqrt{\pi n^3}} + O\left(\frac{1}{n^{5/2}}\right)$
References

1. Gi-Sang Cheon and Louis W. Shapiro, Protected points in ordered trees, Appl Math Lett 21 (2008), 516–520.
2. Keith Copenhaver, k-protected vertices in unlabeled rooted plane trees, Graphs and Combinatorics 33 (2017), no. 2, 347–355.
3. N.G. de Bruijn, D.E. Knuth, and S.E. Rice, The average height of planted plane trees, Graph Theory and Computing (1972), 1522.
4. R.C. Entringer, A. Meir, J.W. Moon, and L.A. Székely, On the wiener index of trees from certain families, Australasian Journal of Combinatorics 10 (1994), 211–224.
5. Philippe Flajolet and Andrew Odlyzko, The average height of binary trees and other simple trees, Journal of Computer and System Sciences 25 (1982), no. 2, 171–213.
6. Philippe Flajolet and Robert Sedgewick, Analytic combinatorics, Cambridge University Press, 2009.
7. Clemens Heuberger and Helmut Prodinger, Protection number in plane trees, Applicable Analysis and Discrete Mathematics 11 (2017), no. 2, 314–326.

E-mail address: keithc@ufl.edu