ENUMERATION OF TREES WITH LOCAL ORIENTATION BY DEGREE SEQUENCES AND REACHABILITY OF VERTICES

BY

NYARIARO ALBERT OLOO P.

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN PURE MATHEMATICS

DEPARTMENT OF PURE AND APPLIED MATHEMATICS

MASENO UNIVERSITY

List of Tables

DECLARATION

This thesis is my own work and has not been presented for a degree award in any other institution.

Signature ____________________ Date ____________________

Nyariaro Albert Oloo P.,
MSC/MAT/00105/2016

This thesis has been submitted for examination with our approval as the university supervisors

Signature ____________________ Date ____________________

Dr. Isaac O. Okoth,
Maseno University.

Signature ____________________ Date ____________________

Dr. Job O. Bonyo,
Maseno University.

ACKNOWLEDGEMENT

First and foremost, my thanks goes to Dr. Isaac O. Okoth for his generosity and kindness in sharing ideas and for guiding me through this research work. At the beginning, the work appeared to be very complicated and I felt like changing my area of research but with his commitment to our meetings, I got the strength to continue with the work. 'Asante Daktari'. I will not forget the efforts and encouragements from Dr. Job Bonyo who advised me to register for my masters in Maseno University, remained very committed in following up how I was fairing on with my studies.

Secondly, I would like to thank my wife Nancy and my son Kelly for their boundless love and moral support. You always ensured that I attended classes and reminded me to read always. You remain the source of my joy 'Kongoi mising.' My sincere thanks also goes to my dad Mr. Charles Oloo, my siblings, my mother in-law and colleagues who kept on encouraging me to finish my work as fast as I could. Mose, Fidel, Meshack and Gori thank you for your time.

Above all, I thank Almighty God for giving me good health throughout this process.

DEDICATION

To my wife Nancy and my son Kelly ...

ABSTRACT

Trees in which edges are oriented from a vertex of lower label towards a vertex of higher label, commonly referred to as locally oriented trees, were introduced by Du and Yin in an attempt to solve a problem conjectured by Ethan Cotterill in his study of secant planes in Algebraic Geometry. Du and Yin, Shin and Zeng, and Stephan Wagner provided proofs for a formula which counts the number of locally oriented trees with a given indegree sequence. Recent studies have concentrated on finding the number of these trees in which both indegree and outdegree sequences are simultaneously given. In this thesis, formulas for the number of locally oriented trees with one source and given outdegree sequences are obtained. Moreover, reachability questions on vertices of locally oriented trees and locally oriented noncrossing trees (first studied by Okoth) have been extensively answered though equivalent results for locally oriented ordered trees had not been obtained. The purpose of this study was to enumerate trees with local orientation by indegree and outdegree sequences as well as reachability of vertices. The specific objectives were; to establish a closed formula for the number of locally oriented trees whose indegree and outdegree sequences are simultaneously given and, to determine formulas counting the number of reachable vertices in labelled ordered trees with local orientation according to path lengths, first children, non-first children, sinks, leaf sinks, non-leaf sinks and left most paths. To achieve the first objective, we used induction approach as well as construction approach to develop recurrence relations. We then used generating functions to find closed formulas. For the second objective, we used construction approach of Seo and Shin, recurrence relations, generating functions and direct proofs. We have obtained closed formulas for reachable vertices in labelled plane trees with respect to: path lengths, sinks, leaf sinks, left most path, first children, non first children and non leaf sinks. The results obtained in this work will add to the already existing literature in this area of research and will also be of importance to computer scientists as most data in computers are stored in form of plane trees.

Table of Contents

Index of Notations

List of Figures

CHAPTER 1

INTRODUCTION

This study is within tree enumeration, a branch of enumerative combinatorics. Enumerative combinatorics is considered by many as the most classical area of combinatorics and it focuses on determining the number of objects satisfying a given property or the number of ways that certain patterns can be formed. Here, our counting objects are labelled trees and labelled ordered trees whose edges are oriented from a vertex of lower label towards a vertex of higher label (which we refer to as *locally oriented trees* and *locally oriented ordered trees* respectively).

1.1 Basic concepts

In this section, we provide definitions and concepts which are well documented in enumerative combinatorics texts, for example, Stanley [23].

1.1.1 Graph theoretic concepts

The following definitions are in [7].

A *graph* G is a pair (V, E) where V is a set of vertices and E is a set of edges. We write $V(G)$ for a set of vertices and $E(G)$ for set of edges of the graph G. Also

$|G| = |V(G)|$ denotes the number of vertices (*order*) and $e(G) = |E(G)|$ denotes the number of edges (*size*). A *simple graph* is a graph with no loops or multiple edges, where a loop is an edge from a vertex into itself and multiple edges are two or more edges joining any two vertices. We say that vertices u and v are *adjacent* in G if they are connected by a common edge and two edges are *incident* if they share a common vertex. The number of vertices in G that are adjacent to a vertex V is the *degree* of V denoted by $deg_G(V)$.

A graph $H =$ *(U,F)* is a *subgraph* of $G =$ *(V,E)* if $U \subseteq V$and $F \subseteq E$. If *U*=*V* then H is called *spanning subgraph* of G. A graph of order n is *complete* if each vertex has degree *n-1*. On the other hand, a graph is *empty* if it has no edges. A *path graph* is a graph that can be drawn so that all its vertices and edges lie on a single straight line. It is also called *line graph*. Whereas a graph of order n in which one of the vertices has degree $n-1$ and the rest of the other vertices have degree 1 is called a *star graph*.

If we assign labels to the vertices of the graph, the resulting graph is called a *labelled graph*. A vertex j is said to be *reachable* from a vertex i if there is a sequence of oriented edges from i to j. A graph in which every vertex is reachable from every other vertex is said to be a *connected* graph. A *tree* is a connected graph without cycles. Therefore, a tree is a subgraph of complete graph. If the number of vertices of the tree equals the number of vertices of the complete graph then the tree is said to be a *spanning tree*. In a tree, a vertex which has degree 1 is called a *leaf*. A labelled graph which is also a tree is called a *labelled tree*. There are n^{n-2} labelled trees on n vertices (Cayley's formula). Figure 1.1 shows a labelled tree on 6 vertices.

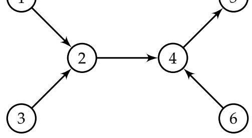


Figure 1.1: Labelled tree

A tree in which a fixed vertex has been chosen is referred to as a *rooted tree* whereas a tree in which a fixed edge has been chosen is referred to as *edge rooted tree*. A *recursive tree* is a tree in which every vertex is reachable from the root. A *forest* is a collection of graphs in which each component is a tree. A rooted forest is a graph whose components are rooted trees. Suppose that T is a rooted tree. If v is a vertex in T other than the root, the *parent* of v is the unique vertex u such that there is a directed edge from u to v. If u is the parent of v then v is the *child* of u.Thus, a rooted tree in which an ordering is specified for the children of each vertex is called an *ordered tree* (or *plane tree*). These trees are counted by $(n-1)^{th}$ Catalan numbers $\frac{1}{n}\binom{2n-2}{n-1}$. Figure 1.2 shows a plane tree on 9 vertices.

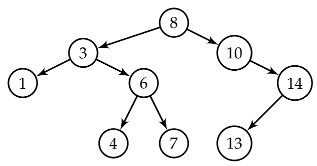


Figure 1.2: Plane tree

A *noncrossing tree* is a tree that can be drawn in the plane with its vertices on the boundary of a circle such that the edges are straight line segments that do not cross. Figure 1.3 is a noncrossing tree of order 6.

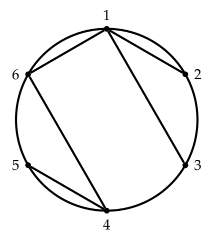
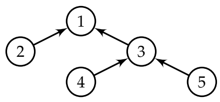


Figure 1.3: Noncrossing tree

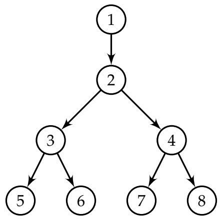
On the other hand, *locally oriented noncrossing trees* are noncrossing trees in which all edges are oriented, from a vertex of lower label towards a vertex with higher label, (See [15]). A *directed graph* (digraph) is a graph in which edges have orientations. The number of edges that are oriented towards a vertex is called *indegree* of the vertex whereas the number of edges that are oriented away from a vertex is called *outdegree*. *Indegree (*resp. *outdegree) sequence* is the ordered sequence of indegree (resp. outdegree) of the vertices of a tree. For instance, if a tree of order n has indegree sequence $\lambda = 0^3 2^2$ then there are 3 vertices with indegree 0 and 2 vertices with indegree 2. A vertex with indegree (resp. outdegree) zero is called *source (sink)* respectively. A *leaf source* is a vertex with indegree 0 and outdegree 1 while a vertex with indegree 1 and outdegree 0 is a *leaf sink*. Two types of orientation have been studied that is: *global orientation* which is a type of orientation in which all the edges in a rooted labelled tree are directed towards/ away from the root, [23]. (See Figure 1.4).



Indegree sequence $\lambda = 0^3 2^2$
Outdegree sequence $\lambda = 0^1 1^4$

Figure 1.4: Tree with global orientation

Local orientation is a type of orientation where each edge is oriented towards the vertex with higher label or from higher label to a lower label. In our case, all the edges will be oriented from a vertex of lower label towards a vertex of higher label, [10]. (See Figure 1.5).



Indegree sequence $\lambda = 0^1 1^7$
Outdegree sequence $\lambda = 0^4 1^1 2^3$
Vertex 1 is a leaf source
Vertices $5, 6, 7, 8$ are leaf sinks

Figure 1.5: Tree with local orientation

1.1.2 Recurrence relations

A *recurrence relation* is an equation in which a term is expressed in terms of preceding terms. A common recurrence relation is the one for Fibonacci numbers which is given as $F_n = F_{n-1} + F_{n-2}$ with initial conditions $F_0 = 0$ and $F_1 = 1$. The following theorem is important in our enumeration:

Theorem 1.1.1. *(see [5]) Suppose that $b_n = Q(n)q^n$, where Q is a polynomial of degree d. If q is not a solution of the characteristic equation, then the solution of the linear recursion*

$$a_n = C_1 a_{n-1} + C_2 a_{n-2} + \cdots + C_r a_{n-r} + b_n$$

is of the form

$$a_n = \sum_{i=1}^{s} P_i(n) q_i^n + P^*(n) q^n$$

where $q_1, q_2, \cdots, q_s$ are (possibly complex) solutions of the characteristic equation

$$q^r = C_1 q^{r-1} + C_2 q^{r-2} + \cdots + C_r$$

and $P_1(n), \cdots, P_s(n), P^*(n)$ are polynomials. The degree of P_i is strictly less than the multiplicity of q_i as a solution of the characteristic equation and the degree of $P^*(n)$ is d.

1.1.3 Generating functions and Functional Equation

Generating functions are one of the most powerful and versatile tools used in enumerative combinatorics [26]. There are two kinds of generating functions. Let $(g_0, g_1 \ldots)$ be a sequence of integers. The ordinary generating function of this sequence is

$$\sum_{i \geq 0} g_i x^i$$

and its exponential generating function is

$$\sum_{i \geq 0} g_i \frac{x^i}{i!}.$$

Ordinary generating functions are usually used to count unlabelled structures. On the other hand, exponential generating functions are used when dealing with labelled structures. Consider for example, the sequence (1, 1, ...). Its ordinary generating function is

$$\sum_{i \geq 0} x^i = \frac{1}{1-x}.$$

while the corresponding exponential generating function is

$$\sum_{i \geq 0} \frac{x^i}{i!} = \exp(x).$$

A *functional equation* is any equation in which the unknown represents a function [22]. Often, the equation relates the value of a function at some point with its values at other points. A functional equation cannot be simply reduced to algebraic equations. An example of a functional equation is the Cauchy functional equation

$$f(x+y) = f(x) + f(y).$$

We shall always denote the coefficient of x^n in a generating function $g(x)$ by $[x^n]g(x)$. Let $f(x)$ be a generating function that satisfies the functional equation

$$f(x) = x\phi(f(x)).$$

We have used the following formula where applicable to extract the coefficient of x^n in the generating function $f(x)$.

Theorem 1.1.2 (Lagrange Inversion Formula)**.** *[See [23] Theorem 5.4.2]. Let $f(x)$ be a generating function that satisfies the functional equation $f(x) = x\phi(f(x))$, where $\phi(0) \neq 0$. We have*

$$[x^n]f(x)^k = \frac{k}{n}[t^{n-k}]\phi(t)^n.$$

For more background on generating functions, we refer to Wilf's *Generatingfunctionology* [26].

1.2 Statement of the problem

The generating function for the number of trees with local orientation such that both indegree and outdegree sequence are simultaneously given already exists in literature. However, the closed formula for such trees is non-existent.

Moreover, the formulas that count the number of reachable vertices in locally oriented trees as well as locally oriented noncrossing trees already exist. For labelled ordered trees with local orientation, an equivalent formula had not been obtained.

1.3 Objectives of the study

The main objective of this study was to enumerate labelled trees with a local orientation by indegree and outdegree sequences as well as reachability of vertices in labelled plane trees.

The specific objectives of the study were:

1. To establish a closed formula for the number of locally oriented trees whose indegree and outdegree sequences are simultaneously given.
2. To determine formulas counting the number of reachable vertices in labelled ordered trees with a local orientation according to path lengths, first children, non-first children, sinks, leaf sinks, non-leaf sinks and left most path.

1.4 Methodology

To achieve our first objective, we have used induction approach as well as construction approach to develop recurrence relations. We have then used known generating functions, where possible, to obtain closed formulas. To establish a closed formula that counts the number of trees with 1 source and ℓ sinks we have constructed a bijection between the set of these trees and the set of recursive trees with ℓ leaves.

For the second objective, we have used construction approach of Seo and Shin [20], generating functions and direct proofs to obtain counting formulas for reachable vertices in labelled ordered trees. Here, Lagrange Inversion Formula, Binomial Theorem and Hockey stick identity have been used extensively.

1.5 Significance of the study

Locally oriented trees have been enumerated by many mathematicians and computer scientists alike, although this enumeration has not been fully exhausted especially on getting closed formulas counting number of trees where either both indegree and outdegree sequences are simultaneously given or where number of sources and sinks are given at the same time. It is therefore of great importance to determine closed formulas for such trees. Also, reachability questions in graphs have been studied for a very long time by mathematicians and computer scientists. This is important in programming. Moreover, most of data in computers are stored as trees (ordered trees). Therefore reachability questions that will be answered in this work will be of significance to the field of computer science.

CHAPTER 2

LITERATURE REVIEW

In this chapter, we review literature related to enumeration of locally oriented trees by indegree and outdegree sequences as well as reachability of vertices.

In 1889, Cayley [2] showed that the number of distinct spanning trees of complete graphs with vertex set $[n] := \{1, ..., n\}$ is given by

$$T_n = n^{n-2}.$$

However, it was noted that an equivalent result was proved earlier in 1860 by Borchdart [1]. This result appeared without proof in an even earlier paper in 1857 by Sylvester [24]. The formula for the number of labelled trees has been rediscovered, conjectured, proved and generalized by many researchers. Some of the proofs include; Kirchhoff's matrix tree theorem [3], Joyal's bijective proof [12], Pitman's double counting argument [16] and Prüfer sequences [18] that yield a bijective proof.

A classical refinement of Cayley's formula for the number of labelled trees with n vertices is obtained by taking vertex degree into account as well, one version can be described by considering an orientation of edges for instances, fix a vertex as a root and assume that all edges are oriented towards the root (global orientation) this gives rise to an indegree sequence $\lambda = 0^{e_0}1^{e_1}2^{e_2}\cdots$. The number of rooted labelled trees with n vertices and λ as the indegree sequence is given by

$$\frac{(n-1)!^2}{e_0(0!)^{e_0}e_1(1!)^{e_1}e_2(2!)^{e_2}\cdots}.$$

(see Stanley [23]). We note that indegree sequence$\lambda = 0^{e_0}1^{e_1}2^{e_2}\cdots$ must satisfy the coherence condition $\sum_i e_i = n$ and $\sum_i ie_i = n-1$.

In 2010, Du and Yin [10] as well as Shin and Zeng [21] using bijective proof showed that this expression also counts the number of labelled trees on n vertices whose indegree sequences is $\lambda = 0^{e_0}1^{e_1}2^{e_2}\cdots$ with respect to local orientation of the vertices. By symmetry it also counts trees with a given outdegree sequence with respect to local orientation. This formula was originally conjectured by Cotterill [6] in the context of algebraic geometry and an extension to rooted forests was proven by Stephan Wagner [25].

Recently, Okoth in his PhD thesis, [14, Theorem 3.2.4] provided a refinement for the above formula by considering a tree of order n with indegree sequence $\lambda = 0^{e_0}1^{e_1}2^{e_2}\cdots$ such that vertex r is a sink of degree d. He obtained the formula as

$$\frac{(r-1)!(n-2)!(n-d-1)!de_d}{(r-d-1)!e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\ldots}$$

Setting $r = n$ and summing over d, we get the formula of Du and Yin [10]. The generating function for the number of trees with a given indegree sequence and outdegree sequence was given by Remmel and Williamson [19] as well as Martin and Reiner [13] as

$$\sum_{T\in Tree(K_n)}\prod_{i=1}^{n} y_i^{indeg_T(i)} x_i^{outdeg_T(i)}$$
$$= x_1y_n\prod_{i=2}^{n-1}((x_1+\cdots+x_i)y_i + x_i(y_{i+1}+\cdots+y_n)).$$

Here, $Tree(Kn)$ is the set of all spanning trees of the complete graph Kn. This generating function is not in the form in which we can use multivariate Lagrange Inversion Formula. There is no known tool for extracting the coefficients, therefore the closed formula does not exist. In this work, we have obtained a recurrence relation satisfied by labelled trees with local orientation in which both indegree sequence and outdegree sequence are given simultaneously as well as explicit formulas for some special cases where indegree sequence is $\lambda = 0^1 1^{n-1}$.

In 1990, William Chen [4] provided a refinement of the Cayley's formula by obtaining a closed formula for number of trees of order n with k sources as

$$\frac{(n-1)!}{k!} \left\{ \begin{matrix} n-1 \\ n-k \end{matrix} \right\}, \tag{2.1}$$

where $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ denotes Sterling numbers of the second kind which count set partitions.

In 2015, Okoth in his PhD thesis, [14, Corollary 3.2.5] provided a formula for number of spanning trees of a complete digraph with local orientation on n vertices having k sources and vertex r as a sink of degree d to be

$$\frac{(r-1)!(n-2)!}{(r-d-1)!(d-1)!k!} \left\{ \begin{matrix} n-d-1 \\ n-k-1 \end{matrix} \right\}.$$

Setting $r = n$ and summing over all $d's$, we rediscover Chen's result given by (2.1). A closed formula for the number of these trees of order n with k sources and ℓ sinks such that $k + \ell = n$ was obtained by Postnikov [17]. The recurrence relation for the number of trees of order n with k sources and ℓ sinks was obtained in Okoth's thesis as a multisum (See Equation (3.3.4) in [14]) and as functional equation (See Theorem 3.3.1 in [14]). This multisum and the functional equation cannot be solved to get a closed formula. In this work, we have determined a closed formula for a specific case when ($k = 1$ or $\ell = 1$).

In 2015, Okoth [14] considered reachability of vertices in locally oriented trees and locally oriented noncrossing trees. He obtained quite a number of statistics of these trees, among them; number of trees in which a given vertex j is reachable from a vertex i, number of reachable vertices, trees with exact number of reachable vertices, number of children of a given vertex, number of recursive trees (a tree in which every vertex is reachable from a given vertex (root)).

Unlabelled ordered trees on n vertices are counted by $(n-1)^{th}$ Catalan numbers

$$C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}.$$

For the labelled ordered trees with n vertices, equivalent formula is

$$n!C_{n-1}, \tag{2.2}$$

(See Seo and Shin [20]). Seo and Shin [20], gave a formula that count the number of forests consisting of k ordered trees on n vertices as

$$f(n,k) = \binom{n}{k} k(n+1)(n+2)\cdots(2n-k-1).$$

The same authors gave the formula for the number of labelled ordered trees in which exactly k vertices are reachable from the root in terms of some trees introduced in the PhD thesis of Drake [8]. For labelled ordered trees with a global orientation, the closed formula that counts the number of all first children at level ℓ, non-first children at level $\ell - 1$, leaves at level ℓ and non-leaves at level $\ell - 1$ reachable from the root is

$$\frac{\ell}{n}\binom{2n}{n+\ell}.$$

(See Eu, Seo and Shin [11] for details). No work has been done for the case of labelled ordered trees with local orientation. Therefore, in this work, we have extended the earlier studies by Okoth [14], Seo and Shin [20] as well as Eu, Seo and Shin [11] to study reachability in locally oriented ordered trees.

CHAPTER 3

ENUMERATION OF TREES WITH LOCAL ORIENTATION

In this chapter, we enumerate locally oriented trees with indegree sequence $0^1 1^{n-1}$ and a given outdegree sequence. Moreover, we obtain a closed formula for the number of trees with one source and a given number of sinks. The results of this chapter are new.

3.1 Enumeration by indegree and outdegree sequences

We begin by proving this elementary but useful fact concerning indegree and outdegree sequences of spanning trees of complete digraphs.

Lemma 3.1.1. *Let G be a complete digraph on n vertices with a local orientation. Also let $0^{f_0}1^{f_1}\cdots$ and $0^{e_0}1^{e_1}\cdots$ be indegree sequences of some spanning trees of G. Then a sequence $0^{f_0}1^{f_1}\cdots$ is an outdegree sequence of a spanning tree with indegree sequence $0^{e_0}1^{e_1}\cdots$ if and only if $e_0 + f_0 \leq n$.*

Proof. $\Longrightarrow$ By contradiction, assume that $e_0 + f_0 > n$. Then $e_0 > n - f_0$. Since $0^{f_0}1^{f_1}\cdots$ is an indegree sequence of some tree then $\sum_{i\geq 0} f_i = n$. Therefore $e_0 > \sum_{i\geq 1} f_i$. This means that the number of vertices with indegree 0 exceeds the total number of vertices with *outdegree* greater than 0, that is, there is at least a vertex with both indegree and outdegree zero. This is impossible in a tree. Thus the assumption is wrong.

$\Longleftarrow$ We shall now assume that $e_0 + f_0 \leq n$ and show that $0^{f_0}1^{f_1}\cdots$ is an outdegree sequence of a spanning tree with indegree sequence $0^{e_0}1^{e_1}\cdots$. Since $0^{e_0}1^{e_1}\cdots$ is an indegree sequence then $\sum_{i\geq 0} e_i = n$. So $e_0 + f_0 \leq n = \sum_{i\geq 0} e_i = \sum_{i\geq 1} e_i + e_0$. It implies that $\sum_{i\geq 1} e_i \geq f_0$. This means that the number of vertices with indegree greater than 0 exceeds or is equal to the number of vertices with outdegree 0, that is, there is no vertex with both indegree and outdegree zero. So, there is no chance of having empty digraphs, and since $0^{f_0}1^{f_1}\cdots$ is an indegree sequence of another tree, then it is an outdegree sequence of a tree with indegree sequence $0^{e_0}1^{e_1}\cdots$. □

We prove the main result of this chapter:

Theorem 3.1.2. *Let G be a complete graph on n vertices. Let $P(n)$ be the number of spanning trees of G with local indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^{f_0}1^{f_1}2^{f_2}\cdots p^{f_p}$ then $P(n)$ satisfies the recurrence relation*

$$P(n) = f_0 P(n-1) + \sum_{i\geq 2}^{p} (f_{i-1}+1) R_i(n-1),$$

where p is the largest integer such that $f_p \neq 0$, and $R_i(n-1)$ is the number of trees of order $n-1$ with indegree sequence $0^1 1^{n-2}$ and outdegree sequence

$0^{f_0-1}1^{f_1}2^{f_2}\cdots(i-1)^{f_{i-1}+1} i^{f_i-1}(i+1)^{f_{i+1}}\cdots.$

Proof. We prove the recurrence by construction. We obtain a tree of order n with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^{f_0}1^{f_1}2^{f_2}\cdots p^{f_p}$ by attaching a vertex of label n to a tree of order $n-1$ with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^{f_0}1^{f_1-1}2^{f_2}\cdots p^{f_p}$. There are two cases to consider:

I. Attachment done at sinks. This gives a total of $f_0 P(n-1)$ trees.

II. Attachment done at a vertex v of outdegree $i-1$.

Consider a tree of order $n-1$ with an indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^{f_0-1}1^{f_1}2^{f_2}\cdots(i-1)^{f_{i-1}+1}i^{f_i-1}(i+1)^{f_{i+1}}\cdots$. This attachment increases the outdegree of vertex v from $i-1$ to i and decreases the number of vertices with outdegree $i-1$ by 1 while the outdegrees of the other vertices remain the same. Note that the new tree has indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^{f_0}1^{f_1}2^{f_2}\cdots p^{f_p}$. The number of these trees are therefore $(f_{i-1}+1)R_i(n-1)$. Thus the total number of these trees is $\sum_{i\geq 2}^{p}(f_{i-1}+1)R_i(n-1)$. This completes the proof. □

Corollary 3.1.3. *Let G be a complete graph on n vertices. Let $P(n)$ be the number of spanning trees of G with local indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^{f_0}1^{f_1}2^{f_2}\cdots p^{f_p}$ where $f_1 = 0$, then $P(n)$ satisfies the following recurrence relation*

$$P(n) = \sum_{i\geq 2}^{p}(f_{i-1}+1)R_i(n-1),$$

where p is the largest integer such that $f_p \neq 0$, and $R_i(n-1)$ is the number of trees of order $n-1$ with indegree sequence $0^1 1^{n-2}$ and outdegree sequence

$0^{f_0-1}1^{f_1}2^{f_2}\cdots(i-1)^{f_{i-1}+1}i^{f_i-1}(i+1)^{f_{i+1}}\cdots$.

Proof. The proof is immediate as no attachment is done at a sink. □

Example 3.1.4. *Consider trees on 7 vertices with indegree sequence $0^1 1^6$ and outdegree sequence $0^4 1^1 2^1 3^1$. These trees can be constructed in two ways:*

(i) Attaching vertex 7 at sink in trees with indegree sequence $0^1 1^5$ and outdegree sequence $0^4 2^1 3^1$. The attachment increases the number of vertices with outdegree 1 by 1 while the number of vertices with outdegree 0 remain the same. Thus the outdegree sequence of the new tree is $0^4 1^1 2^1 3^1$. There are 4 sinks where this attachment can be done. In total there are $4P(6)$ obtained by the attachment if $P(6)$ is the number of trees with indegree sequence $0^1 1^5$ and outdegree sequence $0^4 2^1 3^1$.

(ii) Attaching vertex 7 at a vertex of outdegree 1 in trees with indegrees sequence $0^1 1^5$ and outdegree sequence $0^3 1^2 3^1$. This attachment increases the number of vertices with outdegree 0 by 1 and decreases the number of vertices with outdegree 1 by 1. Thus if $R_1(6)$ is the number of trees with indegree sequence $0^1 1^5$ and outdegree sequence $0^3 1^2 3^1$ then the number of trees with outdegree sequence $0^4 1^1 2^1 3^1$ obtained by this attachment is $2R_1(6)$. Putting everything altogether, there are $4P(6) + 2R(6)$ trees of order 7 with indegree sequence $0^1 1^6$ and outdegree sequence $0^4 1^1 2^1 3^6$.

Corollary 3.1.5. *There is only one tree of order n with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^1 1^{n-1}$ or $0^{n-1}(n-1)^1$.*

Proof. There is only one tree with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^1 1^{n-1}$, that is, a path graph. Also, the only tree with the indegree sequence $0^1 1^{n-1}$and outdegree sequence $0^{n-1}(n-1)^1$ is the star graph. Since these are the only trees with the stated indegree and outdegree sequences, the proof is immediate. □

Corollary 3.1.6. *Let G be a complete digraph on $n \geq 3$ vertices with a local orientation. The number of spanning trees of G with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^2 1^{n-3} 2^1$ is given by $2^{n-1} - n$.*

Proof. Consider a tree with n vertices whose indegree sequence is $0^1 1^{n-1}$ and outdegree sequence is $0^2 1^{n-3} 2^1$. Let the number of these trees of order n be denoted by T_n. These trees can be obtained in two ways. Firstly, from trees on $n-1$ vertices with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^2 1^{n-4} 2^1$ by attaching vertex n at any of the two sinks in the tree hence we have $2T_{n-1}$ such trees. Secondly, the trees can be obtained from line graphs with $n-1$ vertices by gluing vertex n at a source and at a vertex which is neither a source nor a sink. Therefore, gluing can be done at $n-2$ vertices. Thus, the recurrence relation for the number of trees of order n with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^2 1^{n-3} 2^1$ is

$$T_n = 2T_{n-1} + (n-2)L_{n-1}$$

where L_{n-1} is the number of line graphs on $n-1$ vertices, given in Corollary 3.1.5 as 1.

Hence,

$$T_n = 2T_{n-1} + n - 2.$$

The initial condition is $T_3 = 1$. In the remaining part of the proof, we obtain a closed formula from this recurrence relation as:

Let $T_n = r^n$. Then, solving the characteristic equation $T_n = 2T_{n-1}$, we obtain $r = 2$. From Theorem 1.1.1, we have

$$T_n = A2^n + (Bn + C)1^n \tag{3.1}$$

Replacing n with $n-1$ we get,

$$T_{n-1} = A2^{n-1} + (Bn - B + C)1^{n-1}$$

Multiplying by 2 we obtain

$$2T_{n-1} = A2^n + (2Bn - 2B + 2C)1^{n-1}$$

Comparing the coefficients and solving for the unknowns we obtain, $A = \frac{1}{2}$, $B = -1$ and $C = 0$. Thus substituting the unknowns in Equation (3.1) we obtain,

$$T_n = 2^{n-1} - n$$

Thus the proof. □

Corollary 3.1.7. *Let G be a complete digraph on $n \geq 4$ vertices with a local orientation. The number of spanning trees of G with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^3 1^{n-4} 3^1$ is given by*

$$\frac{1}{4}\left[3^{n-1} - 2^{n+1} + 2n + 1\right].$$

Proof. We induct on n.

Base case, $n = 4$: The number of spanning trees with indegree sequence $0^1 1^3$ and outdegree sequence $0^3 3^1$ is 1. (See Corollary 3.1.5). So base case holds.

Induction step: We assume that the number of spanning trees of a complete digraph on $n-1$ vertices with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^3 1^{n-4} 3^1$ is

$$T_{n-1} = \frac{1}{4}\left(3^{n-2} - 2^n + 2n - 1\right),$$

and prove that it holds for a graph with n vertices. We construct a tree with n vertices by gluing an edge and a vertex (with label n) to the existing tree with $n-1$ vertices. In the new tree, the number of vertices with indegree 1 increases by 1. Since we have a local orientation and the vertex to be added has the highest label, then it can be glued at a vertex of outdegree zero giving a total of

$$\frac{3}{4}\left(3^{n-2} - 2^n + 2n - 1\right)$$

new trees.

The remaining spanning trees on n vertices with the given indegree and outdegree sequences can be constructed as follows: Consider a digraph on $n-1$ vertices with a local orientation such that its indegree sequence is $0^1 1^{n-2}$ and outdegree sequence is $0^2 1^{n-4} 2^1$. Now, glue vertex n and its incident edge at each vertex with outdegree 2. By Corollary 3.1.6, we obtain $2^{n-2} - n + 1$ such trees. The total number of spanning trees on n vertices is therefore

$$\begin{aligned}
T_n &= \frac{3}{4}\left(3^{n-2} - 2^n + 2n - 1\right) + 2^{n-2} - n + 1 \\
&= \frac{1}{4}\left(3^{n-1} - 3\cdot 2^n + 6n - 3 + 4\cdot 2^{n-2} - 4n + 4\right) \\
&= \frac{1}{4}\left(3^{n-1} - 3\cdot 2^n + 6n - 3 + 2^n - 4n + 4\right) \\
&= \frac{1}{4}\left(3^{n-1} - 2^{n+1} + 2n + 1\right).
\end{aligned}$$

Hence the proof. □

Corollary 3.1.8. *Let G be a complete digraph on $n \geq 5$ vertices with a local orientation. The number of spanning trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^3 1^{n-5} 2^2$ is given by:*

$$\frac{1}{4}\left[3^n - (n-1)2^{n+1} + 2n^2 - 4n - 1\right].$$

Proof. We construct trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^3 1^{n-5} 2^2$ by gluing the n^{th} vertex to a tree on $n-1$ vertices. The required indegree and outdegree sequences are achieved by attaching the extra vertex to a vertex of outdegree 1 in a tree with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^2 1^{n-4} 2^1$. Hence by Corollary 3.1.6, we obtain

$$(n-4)T_{n-1}$$

new trees where

$$T_{n-1} = 2^{n-2} - (n-1).$$

The remaining spanning trees can be obtained by gluing a vertex of label n to a vertex of outdegree 0 in a tree of indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^3 1^{n-6} 2^2$. Therefore, we have $3P_{n-1}$ new trees. Since these trees have outdegree sequence $0^3 1^{n-5} 2^2$ then if $n \leq 4$, we obtain a negative number of trees with outdegree 1. This is impossible and we impose $P_4 = 0$. The total number of spanning trees is therefore given recursively as:

$$P_n = 3P_{n-1} + (n-4)T_{n-1} \text{ for } n \geq 5, \tag{3.2}$$

where $P_4 = 0$ and $T_{n-1} = 2^{n-2} - n + 1$.

We now find explicit formula from recursion (3.2):

Let $P_n = 3P_{n-1} + (n-4)T_{n-1}$ and $T_{n-1} = 2^{n-2} - n + 1$. We have,

$$\begin{aligned} P_n &= 3P_{n-1} + n.2^{n-2} - n^2 + n - 4.2^{n-2} + 4n - 4 \\ &= 3P_{n-1} + \left(\frac{n}{2} - 2\right)2^{n-1} + (-n^2 + 5n - 4)1^{n-1}. \end{aligned}$$

Now, let $P_n = r^n$. Then, solving the characteristic equation $P_n = 3P_{n-1}$, we obtain $r = 3$. From Theorem 1.1.1, we have

$$P_n = A3^n + (Bn + C)2^n + (Dn^2 + En + F)1^n. \tag{3.3}$$

Replacing n with $n-1$ we get,

$$P_{n-1} = A3^{n-1} + (Bn - B + C)2^{n-1} + (D(n^2 - 2n + 1) + E(n-1) + F)1^{n-1}.$$

Multiply by 3 we obtain,

$$\begin{aligned}3P_{n-1} &= A3^n + (3Bn - 3B + 3C)2^{n-1} + (3Dn^2 - 6Dn + 3D + 3En - 3E + 3F)\\ &= A3^n + \left(\frac{3Bn - 3B + 3C}{2}\right) 2^n + (3Dn^2 - 6Dn + 3D + 3En - 3E + 3F).\end{aligned}$$

Comparing the coefficients and solving for the unknowns we obtain, $A = \frac{1}{4}$, $B = -\frac{1}{2}$, $C = \frac{1}{2}$, $D = \frac{1}{2}$, $E = -1$ and $F = -\frac{1}{4}$. Thus substituting the unknowns in Equation (3.3) we obtain,

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

This completes the proof. $\square$

Corollary 3.1.9. *Let G be a complete digraph on $n \geq 5$ vertices with a local orientation. The number of spanning trees of G with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^4 1^{n-5} 4^1$ is given by:*

$$\frac{1}{72}\left[4^n - 2 \cdot 3^{n+1} + 9 \cdot 2^{n+1} - 12n - 10\right].$$

Proof. We construct trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^4 1^{n-5} 4^1$ by gluing the n^{th} vertex to a tree on $n-1$ vertices. The required indegree and outdegree sequences are obtained firstly, by attaching the vertex n to a vertex of outdegree 3 in a tree of indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^3 1^{n-5} 3^1$. Hence by Corollary 3.1.7, we obtain

$$\frac{1}{4}\left(3^{n-2} - 2^n + 2n - 1\right) = T_{n-1}$$

new trees.

The remaining trees can be obtained by gluing vertex n to a vertex of outdegree 0 in a tree of indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^4 1^{n-6} 4^1$. Therefore, we have $4P_{n-1}$ new trees. The total number of the trees is thus

$$P_n = 4P_{n-1} + T_{n-1} \tag{3.4}$$

for $n \geq 5$, where $P_4 = 0$ and $T_{n-1} = \frac{1}{4}\left(3^{n-2} - 2^n + 2n - 1\right)$. The desired result follows by finding the explicit formula of Equation 3.4; We have,

$$\begin{aligned} P_n &= 4P_{n-1} + \frac{1}{4}\left(3^{n-2} - 2^n + 2n - 1\right) \\ &= 4P_{n-1} + \frac{1}{4}\cdot 3^{n-2} - \frac{1}{4}\cdot 2^n + \frac{1}{2}n - \frac{1}{4}. \end{aligned}$$

Let $P_n = r^n$. Then, solving the characteristic equation, $P_n = 4P_{n-1}$, we obtain $r = 4$. Hence we rewrite P_n as;

$$P_n = A4^n + B3^n + C2^n + Dn + E. \tag{3.5}$$

(See Theorem 1.1.1). Replacing n with $n-1$ we obtain;

$$P_{n-1} = A4^{n-1} + B3^{n-1} + C2^{n-1} + D(n-1) + E.$$

Multiplying by 4 we have,

$$\begin{aligned} 4P_{n-1} &= A4^n + 4B3^{n-1} + 4C2^{n-1} + 4Dn - 4D + 4E \\ &= A4^n + 4B3^{n-1} + 2C2^n + 4Dn - 4D + 4E. \end{aligned}$$

Comparing the coefficients and solving for the unknowns we obtain, $A = \frac{1}{72}$, $B = -\frac{1}{12}$, $C = \frac{1}{4}$, $D = -\frac{1}{6}$ and $E = -\frac{5}{36}$. Now, substituting these values in Equation (3.5) we get

$$P_n = \frac{1}{72}\left(4^n - 2\cdot 3^{n+1} + 9\cdot 2^{n+1} - 12n - 10\right).$$

This is the desired formula. □

Corollary 3.1.10. *Let G be a complete digraph on $n \geq 5$ vertices with a local orientation. The number of spanning trees of G with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^4 1^{n-6} 2^1 3^1$ is given by:*

$$\frac{1}{36}\left[4^{n+1} - (n+5)\cdot 3^{n+1} + 27(n-1)\cdot 2^n - 18n^2 + 33n + 23\right].$$

Proof. We construct trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^4 1^{n-6} 2^1 3^1$ by gluing vertex n to a tree on $n-1$ vertices. The required indegree and outdegree sequence are achieved by attaching vertex to a vertex n of outdegree 1 in a tree with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^3 1^{n-5} 3^1$. Hence by Corollary 3.1.7, we obtain that there are

$$(n-5) \cdot \frac{1}{4}\left(3^{n-2} - 2^n + 2n - 1\right) = (n-5) \cdot T_{n-1}$$

new trees, obtained by this procedure.

The required indegree and outdegree sequences can also be achieved by attaching vertex n to a vertex of outdegree 2 in a tree with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^3 1^{n-5} 2^2$. So by Corollary 3.1.8, we obtain $2 \cdot S_{n-1}$ new trees.

Lastly, the remaining spanning trees are obtained by gluing vertex n to a vertex of outdegree 0 in a tree with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^4 1^{n-7} 2^1 3^1$. Here, we get $4P_{n-1}$ new trees. The total number of spanning trees is thus

$$P_n = 4P_{n-1} + (n-5)T_{n-1} + 2S_{n-1} \tag{3.6}$$

for $n \geq 6$, where $P_5 = 0$ and, T_n and S_n are the number of trees with outdegree sequence $0^3 1^{n-4} 3^1$ and $0^3 1^{n-5} 2^2$ respectively.

We solve the recursion equation (3.6) above to obtain the desired closed formula.

□

Corollary 3.1.11. *Let G be a complete digraph on $n \geq 7$ vertices with a local orientation. The number of spanning trees of G with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^4 1^{n-7} 2^3$ is given by:*

$$P_n = \frac{1}{24}\left[3 \cdot 4^n - 6(n-1) \cdot 3^n + 6(n^2 - 4n + 2) \cdot 2^n - 4n^3 + 24n^2 - 26n - 12\right]. \tag{3.7}$$

Proof. We construct trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^4 1^{n-7} 2^3$ by gluing vertex n to a tree on $n-1$ vertices. The required indegree and outdegree sequences are achieved by attaching vertex n to a vertex of outdegree

1 in a tree of indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^3 1^{n-6} 2^2$. By Corollary 3.1.8 we obtain that there are $(n-6)T_{n-1}$ such trees. Here T_n is the number of trees with outdegree sequence $0^3 1^{n-5} 2^2$.

The other trees are arrived at by gluing vertex n to a vertex of outdegree 0 in a tree with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^4 1^{n-8} 2^3$. We have $4P_{n-1}$ such new trees. In total the number of spanning trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^4 1^{n-7} 2^3$ is given by

$$P_n = 4P_{n-1} + (n-6)T_{n-1} \text{ for } n \geq 7, \tag{3.8}$$

where $P_5 = 0$ and, T_n is the number of trees with outdegree sequence $0^3 1^{n-5} 2^2$.

By solving Equation (3.8), we obtain the desired result. □

Corollary 3.1.12. *Let G be a complete digraph on $n \geq 6$ vertices with a local orientation. The number of spanning trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^5 1^{n-6} 5^1$ is given by:*

$$\frac{1}{288}\left[3 \cdot 5^{n-1} - 4^{n+1} + 4 \cdot 3^{n+1} - 3 \cdot 2^{n+2} + 12n + 13\right].$$

Proof. The proof follows by construction: We construct trees with indegree sequence $0^1 1^{n-1}$ and outdegree sequence $0^5 1^{n-6} 5^1$ by gluing a vertex n to a spanning tree on $n-1$ vertices. Here attachment of a vertex n is done at a vertex of outdegree 4 in a tree with $n-1$ vertices and indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^4 1^{n-6} 4^1$. By Corollary 3.1.9, we obtain T_{n-1} trees by this construction.

The remaining trees are obtained by gluing vertex n to a vertex of outdegree 0 in a tree with indegree sequence $0^1 1^{n-2}$ and outdegree sequence $0^5 1^{n-7} 5^1$. Thus, we have $5P_{n-1}$ such new trees. Thus the total number of spanning trees is therefore given by:

$$P_n = 5P_{n-1} + T_{n-1} \text{ for } n \geq 6, \tag{3.9}$$

where $P_5 = 0$ and, T_n is the number of trees with outdegree sequence $0^4 1^{n-5} 4^1$.

Here the initial condition is $P_5 = 0$, since there is no such tree of order 5. We obtain the explicit formula in Corollary 3.1.12 by solving the recursion 3.9. $\square$

3.2 Enumeration by sources and sinks

We prove the following formula which also counts the number of leaves in Cayley trees proved by Drmota in [9].

Proposition 3.2.1. *Let G be a complete digraph on n vertices with a local orientation. The number of spanning trees of G having a single source and ℓ sinks is given by:*

$$\sum_{i=0}^{\ell-1}(-1)^i \binom{n}{i} (\ell - i)^{n-1}. \tag{3.10}$$

Proof. Recursive trees of order n with ℓ leaves are counted by Equation (3.10) (See [9]). We construct a bijection between the set of the recursive trees and the set of spanning trees on n vertices with one source and ℓ sinks. Consider a tree of order n with a single source and ℓ sinks. We take vertex 1 (single source) as the root. Then the internal vertices will neither be sources nor sinks. Then the leaves of the resultant tree are all sinks. This is a recursive tree with ℓ leaves. The process is reversible. Thus the proof. $\square$

Corollary 3.2.2. *Let G be a complete digraph on n vertices with a local orientation. The number of trees on these vertices having a single sink and ℓ sources is given by Equation (3.10).*

Proof. By reversing the orientation of the edges in the graph G we obtain the condition in Proposition 3.2.1 and the proof follows. $\square$

CHAPTER 4

REACHABILITY IN LABELLED ORDERED TREES WITH LOCAL ORIENTATION

4.1 Introduction

In this chapter, we present results on reachability of vertices in labelled ordered trees with local orientation. We examine the number of reachable vertices from a given root i. We also determine a formula for the number of labelled ordered trees on n vertices such that exactly k vertices are reachable from i. In some instances, we look for asymptotic results. In a nutshell, we have enumerated the number of reachable vertices in labelled plane trees by path lengths, first children, left most path, non first children, leaf sinks and non leaf sinks. We considered trees having their edges oriented from a vertex of lower label towards a vertex of higher label (local orientation). This orientation was introduced by Du and Yin in [10].

We recall from Chapter 1 that a vertex j is *reachable* from a vertex i if there is a sequence of oriented edges (paths) from vertex i to vertex j, and a path is of length ℓ if there are ℓ edges on the path. Here, *degree* of a vertex is the number of edges that come out of a vertex if the edges are oriented away from the root.

A vertex in which there is no edge that is oriented away from it is called a *sink* whereas a *leaf sink* is a vertex with only one edge oriented towards it but no edge oriented away from it. The vertices with the same parent are called *siblings*. Since the siblings are linearly ordered, they are always drawn in a left-right pattern where the leftmost sibling is referred to as *first child.* At a level ℓ the left most child is the *eldest child*. A *left most path* refers to a sequence of edges joining eldest children at each level in a plane tree.

We shall use this theorem in our work.

Theorem 4.1.1 (Binomial Theorem)**.** *For any integer* $n \geq 0$, $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

Other useful identities used in this work are; Hockey stick identity,

$$\sum_{k=n}^{m} \binom{k}{n} = \binom{m+1}{n+1}, \tag{4.1}$$

and

$$\sum_{k\geq 0} \binom{-m}{k} (-1)^k = \sum_{k\geq 0} \binom{m+k-1}{k}. \tag{4.2}$$

For the rest of the chapter, we shall refer to labelled ordered trees as trees.

4.2 Enumeration by path lengths

In this section, we enumerate trees by path length which is simply the number of steps (edges) between the vertices.

Theorem 4.2.1. *The number of trees on n vertices rooted at vertex i such that vertex j of degree d is reachable from the root in ℓ steps is given by*

$$\frac{2\ell+d}{n+\ell-1}(n-\ell-1)! \binom{j-i-1}{\ell-1} \binom{2n-d-3}{n+\ell-2}.$$

Proof. Let $P(x)$ be the generating function for plane trees where x marks the number of non-root vertices. Consider a plane tree rooted at vertex i such that there is a path of length ℓ starting at vertex i and ending at vertex $i+\ell$ of degree d. This path decomposes the tree into left and right plane subtrees upto length ℓ hence we have $(P(x)xP(x))^{\ell}$ as the generating function for the number of the trees with a path of length ℓ starting at i and ending at $i+\ell$. (See Figure 4.1). Vertex $i+\ell$ is joined to d other vertices which are connected to other plane trees hence we have $x(xP(x))^{d}$, to represent the vertex and the connected vertices. Putting everything together, we obtain $(P(x)xP(x))^{\ell}x(xP(x))^{d} = x^{\ell+d+1}P(x)^{2\ell+d}$ as the generating function of the unlabelled plane tree rooted at vertex i with a path of length ℓ starting at i and ending at vertex $i+\ell$ of degree d. This is represented pictorially as:

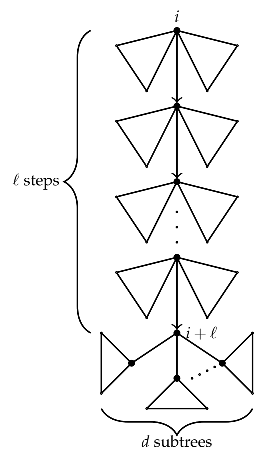


Figure 4.1: Unlabelled ordered tree with path length ℓ.

A plane tree can have any number of children and are thus represented symbolically as

$$P(x) = 1 + xP(x) + x^2P(x)^2 + \cdots .$$

(A plane tree consists of a root to which a sequence, possibly empty, of plane trees is attached.) We have,

$$\begin{aligned}P(x) &= 1 + xP(x)\,(1 + xP(x) + \cdots)\\ &= 1 + xP(x)P(x)\\ &= 1 + xP(x)^2.\end{aligned}$$

Evidently,

$$P(x) = \frac{1}{1 - xP(x)}.$$

This is the generating function for unlabelled plane trees. The generating function $P(x)$, is not in a form that we can use Lagrange inversion formula (1.1.2). Thus, we let $xP(x) = F(x)$ so that

$$P(x) = \frac{F(x)}{x} = \frac{1}{1 - F(x)}.$$

Precisely,

$$F(x) = \frac{x}{1 - F(x)}.$$

Applying Lagrange Inversion Formula (Theorem 1.1.2), we get

$$\begin{aligned}
[x^n]x^{\ell+d+1}P(x)^{2\ell+d} &= [x^n]x^{-\ell+1}F(x)^{2\ell+d} \\
&= [x^{n+\ell-1}]F(x)^{2\ell+d} \\
&= \frac{2\ell+d}{n+\ell-1}[t^{n-\ell-d-1}]\left((1-t)^{-1}\right)^{n+\ell-1} \\
&= \frac{2\ell+d}{n+\ell-1}[t^{n-\ell-d-1}]\,(1-t)^{-(n+\ell-1)}.
\end{aligned}$$

By Binomial Theorem (Theorem 4.1.1), we obtain

$$[x^n]x^{\ell+d+1}P(x)^{2\ell+d} = \frac{2\ell+d}{n+\ell-1}[t^{n-\ell-d-1}]\sum_{i\geq 0}\binom{-(n+\ell-1)}{i}(-t)^i.$$

Hence by Equation (4.2), we arrive at

$$\begin{aligned}
[x^n]x^{\ell+d+1}P(x)^{2\ell+d} &= \frac{2\ell+d}{n+\ell-1}[t^{n-\ell-d-1}]\sum_{i\geq 0}\binom{n+\ell+i-2}{i}t^i \\
&= \frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n-\ell-d-1} \\
&= \frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}.
\end{aligned}$$

This formula counts the number of unlabelled plane trees in which vertex $i+\ell$ of degree d is reachable from the root in ℓ steps. The number of ways of choosing labels for the vertices on a path of length ℓ from vertex i to vertex j is $\binom{j-i-1}{\ell-1}$.

So once the $\ell+1$ vertices on the path are labelled, there are $(n-\ell-1)!$ choices for labelling the remaining vertices. Thus, the number of plane trees in which vertex j of degree d is reachable from root i in ℓ steps is given by

$$(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}\binom{j-i-1}{\ell-1}\binom{2n-d-3}{n+\ell-2}.$$

This completes the proof. $\square$

Setting $\ell=0$ in the just proved theorem, we get that there are

$$(n-1)!\frac{d}{n-1}\binom{2n-d-3}{n-2}=d(n-2)!\binom{2n-d-3}{n-2}$$

trees in which the root has degree d.

Corollary 4.2.2. *The total number of trees on n vertices rooted at vertex i such that vertex j is reachable from the root in ℓ steps is given by:*

$$(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{j-i-1}{\ell-1}\binom{2n-1}{n+\ell}. \tag{4.3}$$

Proof. The result follows by summing over all d in Theorem 4.2.1. Consider the formula

$$\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2},$$

then, we obtain a telescoping form of the formula as

$$\begin{aligned}\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2} &= \frac{(n+\ell-1)-(n-\ell-d-1)}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}\\ &= \frac{n+\ell-1}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}-\frac{n-\ell-d-1}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}\\ &= \binom{2n-d-3}{n+\ell-2}-\frac{n-\ell-d-1}{n+\ell-1}\binom{2n-d-3}{n-\ell-d-1}\\ &= \binom{2n-d-3}{n+\ell-2}-\frac{n-\ell-d-1}{n+\ell-1}\frac{2n-d-3}{n-\ell-d-1}\binom{2n-d-4}{n-\ell-d-2}\end{aligned}$$

on further simplification we obtain,

$$\begin{aligned}\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2} &= \binom{2n-d-3}{n+\ell-2} - \frac{2n-d-3}{n+\ell-1}\binom{2n-d-4}{n+\ell-2} \\ &= \binom{2n-d-3}{n+\ell-2} - \binom{2n-d-3}{n+\ell-1}.\end{aligned}$$

Therefore summing over all d we get

$$\sum_{d=0}^{n-\ell-1}\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2} = \sum_{d=0}^{n-\ell-1}\binom{2n-d-3}{n+\ell-2} - \sum_{d=0}^{n-\ell-2}\binom{2n-d-3}{n+\ell-1}.$$

Setting $k = 2n-d-3$, we obtain

$$\begin{aligned}\sum_{d=0}^{n-\ell-1}\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2} &= \sum_{k=n+\ell-2}^{2n-3}\binom{k}{n+\ell-2} - \sum_{k=n+\ell-1}^{2n-3}\binom{k}{n+\ell-1} \\ &= \binom{2n-2}{n+\ell-1} - \binom{2n-2}{n+\ell}.\end{aligned}$$

Now, simplifying the right hand side, we have

$$\begin{aligned}\sum_{d=0}^{n-\ell-1}\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2} &= \frac{(2n-2)!}{(n+\ell-1)!(n-\ell-1)!} - \frac{(2n-2)!}{(n+\ell)!(n-\ell-2)!} \\ &= \frac{(2n-2)!}{(n+\ell-1)!(n-\ell-2)!}\left(\frac{1}{(n-\ell-1)} - \frac{1}{(n+\ell)}\right) \\ &= \frac{(2n-2)!}{(n+\ell-1)!(n-\ell-2)!}\frac{(n+\ell)-(n-\ell-1)}{(n+\ell)(n-\ell-1)} \\ &= \frac{(2\ell+1)(2n-2)!}{(n+\ell)!(n-\ell-1)!} \\ &= \frac{(2\ell+1)(2n-1)!}{(2n-1)(n+\ell)!(n-\ell-1)!} \\ &= \frac{2\ell+1}{2n-1}\binom{2n-1}{n+\ell}.\end{aligned}$$

Hence,

$$\begin{aligned}&(n-\ell-1)!\binom{j-i-1}{\ell-1}\sum_{d=0}^{n-\ell-1}\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2} \\ &\qquad = (n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{j-i-1}{\ell-1}\binom{2n-1}{n+\ell}.\end{aligned}$$

This is the desired formula. □

Corollary 4.2.3. *The number of vertices in trees of order n that are reachable from root i in ℓ steps is given by*

$$(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{n-i}{\ell}\binom{2n-1}{n+\ell}. \tag{4.4}$$

Proof. We obtain the desired result by summing over all j in Equation (4.3), i.e.,

$$(n-\ell-1)!\frac{2\ell+1}{2n-1}\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}\binom{2n-1}{n+\ell},$$

such trees. Again, we simplify the sum. Let $k = j-i-1$, so that

$$\begin{aligned}(n-\ell-1)!\frac{2\ell+1}{2n-1}&\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}\binom{2n-1}{n+\ell}\\ &= (n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{2n-1}{n+\ell}\sum_{k=\ell-1}^{n-i-1}\binom{k}{\ell-1}.\end{aligned}$$

By hockey stick identity (4.1), we get the required formula as

$$\binom{n-i}{\ell}(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{2n-1}{n+\ell}.$$

Thus the proof. □

Corollary 4.2.4. *There are a total of*

$$(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{n}{\ell+1}\binom{2n-1}{n+\ell} \tag{4.5}$$

vertices that are reachable from the root in ℓ steps, in trees with n vertices.

Proof. We get the result by summing over all i in Equation (4.4), that is,

$$(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{2n-1}{n+\ell}\sum_{i=1}^{n-\ell}\binom{n-i}{\ell}.$$

We set $k = n-i$ to obtain

$$(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{2n-1}{n+\ell}\sum_{k=\ell}^{n-1}\binom{k}{\ell}.$$

Applying hockey stick identity (4.1), the result is immediate. □

Now, summing over all ℓ in Equation (4.5) we obtain

Corollary 4.2.5. *The total number of vertices in trees on n vertices that are reachable from the root is given by*

$$\frac{n!}{2n-1}\sum_{\ell=0}^{n-1}\frac{2\ell+1}{(\ell+1)!}\binom{2n-1}{n+\ell}.$$

Corollary 4.2.6. *The average number of vertices that are reachable from the root in ℓ steps in a random tree is $\frac{2\ell+1}{(\ell+1)!}$.*

Proof. Dividing the total number of vertices that are reachable from the root in ℓ steps in plane trees (See Equation (4.5)) by the total number of labelled plane trees given by Equation (2.2), we get

$$\frac{\binom{n}{\ell+1}(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{2n-1}{n-\ell-1}}{n!\frac{1}{n}\binom{2n-2}{n-1}}$$

as the average number of vertices that are reachable in ℓ steps from the root in trees with n vertices. We simplify the average to get

$$\frac{\binom{n}{\ell+1}(n-\ell-1)!\frac{2\ell+1}{2n-1}\binom{2n-1}{n-\ell-1}}{n!\frac{1}{n}\binom{2n-2}{n-1}} = \frac{2\ell+1}{(\ell+1)!}\frac{(n-1)!n!}{(n+\ell)!(n-\ell-1)!}.$$

Now, taking limits as $n \to \infty$, we get

$$\begin{aligned}\lim_{n\to\infty}\frac{2\ell+1}{(\ell+1)!}\frac{(n-1)!n!}{(n+\ell)!(n-\ell-1)!} &= \frac{2\ell+1}{(\ell+1)!}\lim_{n\to\infty}\frac{(n-1)!n!}{(n+\ell)!(n-\ell-1)!}\\ &= \frac{2\ell+1}{(\ell+1)!}\lim_{n\to\infty}\frac{(n-1)(n-2)(n-3)\cdots(n-\ell)}{(n+\ell)(n+\ell-1)\cdots(n+1)}\\ &= \frac{2\ell+1}{(\ell+1)!}\lim_{n\to\infty}\left(\frac{n^\ell+\cdots}{n^\ell+\cdots}\right)\\ &= \frac{2\ell+1}{(\ell+1)!}.\end{aligned}$$

Hence the desired formula. □

Corollary 4.2.7. *The number of trees of order n in which there is a path of length ℓ starting from the root and ending at a vertex of degree d is given by*

$$(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}\binom{n}{\ell+1}\binom{2n-d-3}{n+\ell-2}. \tag{4.6}$$

Proof. We obtain the result by summing over all i and j in Theorem 4.2.1. That is, the formula is

$$(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}\sum_{i=1}^{n-\ell}\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}\binom{2n-d-3}{n+\ell-2},$$

which we now simplify. Setting $k=j-i-1$, we get

$$(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}\sum_{i=1}^{n-\ell}\sum_{k=\ell-1}^{n-i-1}\binom{k}{\ell-1}.$$

By (4.1), we obtain

$$(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}\sum_{i=1}^{n-\ell}\binom{n-i}{\ell}.$$

Let $m=n-i$ so that

$$\begin{aligned}(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}&\sum_{i=1}^{n-\ell}\binom{n-i}{\ell}\binom{2n-d-3}{n+\ell-2}\\&=(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}\sum_{m=\ell}^{n-1}\binom{m}{\ell}.\end{aligned}$$

Applying hockey stick identity again, we get

$$\begin{aligned}(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}&\binom{2n-d-3}{n+\ell-2}\sum_{i=1}^{n-\ell}\binom{n-i}{\ell}\\&=(n-\ell-1)!\frac{2\ell+d}{n+\ell-1}\binom{n}{\ell+1}\binom{2n-d-3}{n+\ell-2}.\end{aligned}$$

Hence the result. □

Setting $\ell=0$ in Equation (4.6), we get

$$\frac{n!d}{n-1}\binom{2n-d-3}{n-2},$$

as the formula which counts the number of trees on n vertices such that the root is of degree d. Also setting $\ell = 1$ in Equation (4.6), we obtain

$$\frac{(d+2)(n-1)!}{2}\binom{2n-d-3}{n-1}.$$

This formula counts the total number of children of degree d, in all trees of order n.

4.3 Enumeration by sinks and leaf sinks

We recall from Section 4.1 that a *sink* is a vertex with outdegree 0 while a *leaf sink* is a vertex with indegree 1 and outdegree 0. In this section, we enumerate trees with respect to sinks and leaf sinks.

Proposition 4.3.1. *The number of trees of order n in which a vertex j is a sink of degree d reachable from a vertex i in ℓ steps is*

$$(n-\ell-d-1)!\frac{2\ell+d}{n+\ell-1}\binom{j-i-1}{\ell-1}\binom{j-\ell-1}{d}\binom{2n-d-3}{n+\ell-2}. \tag{4.7}$$

Proof. From the proof of Theorem 4.2.1, it follows that there are

$$\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}$$

unlabelled trees with a path of length ℓ starting at the root such that the terminating vertex has degree d. Lets consider a path of length ℓ starting at root i and ending at vertex j. There are $\binom{j-i-1}{\ell-1}$ choices for labelling vertices on the paths. Since vertex j is a sink of degree d, the labels of the d vertices must be less than j. Thus there are $\binom{j-\ell-1}{d}$ choices for the labels. Once the $\ell+1$ vertices on the path and the d children of j are labelled, there are $(n-\ell-d-1)!$ choices for the other labels in the tree. Collecting everything, we arrive at the required formula. □

Corollary 4.3.2. *The total number of sinks of degree d that are reachable from vertex i in ℓ steps in trees of order n is given by*

$$(n-\ell-d-1)!\frac{2\ell+d}{n+\ell-1}\binom{2n-d-3}{n+\ell-2}\sum_{j=\ell+i}^{n}\binom{j-i-1}{\ell-1}\binom{j-\ell-1}{d}.$$

Proof. We obtain the formula by summing over all j in Equation (4.7). ☐

Corollary 4.3.3. *The total number of trees with n vertices such that root i is a sink of degree d is given by:*

$$(n-d-1)!\frac{d}{n-1}\binom{i-1}{d}\binom{2n-d-3}{n-2}. \tag{4.8}$$

Proof. The result follows by setting $\ell = 0$ and $j = i$ in Equation (4.7). ☐

Corollary 4.3.4. *The number of children of the root labelled j having degree d in trees on n vertices is given by:*

$$(n-d-2)!\frac{2+d}{n}\binom{j-2}{d}\binom{2n-d-3}{n-1}. \tag{4.9}$$

Proof. The result is immediate by setting $\ell = 1$ in Equation (4.7). ☐

Summing over all j in Equation (4.9), we obtain the number of children of the root which are sinks of degree d, in trees of order n.

Corollary 4.3.5. *The total number of trees of order n with root sinks of degree d is given by:*

$$(n-d-1)!\frac{d}{n-1}\binom{n}{d+1}\binom{2n-d-3}{n-2}. \tag{4.10}$$

Proof. By summing over all i in Equation (4.8), we obtain

$$(n-d-1)!\frac{d}{n-1}\binom{2n-d-3}{n-2}\sum_{i=d+1}^{n}\binom{i-1}{d}.$$

Now, taking $k = i-1$, we get

$$\begin{aligned}(n-d-1)!\frac{d}{n-1}&\binom{2n-d-3}{n-2}\sum_{i=d+1}^{n}\binom{i-1}{d}\\ &= (n-d-1)!\frac{d}{n-1}\binom{2n-d-3}{n-2}\sum_{k=d}^{n-1}\binom{k}{d}.\end{aligned}$$

Thus, by identity (4.1), we get the desired result.

☐

Corollary 4.3.6. *The average number of root sinks of degree d in a random tree of order n is*

$$\frac{d}{2^{d+1}(d+1)!}. \tag{4.11}$$

Proof. Diving the total number of labelled plane trees of order n with root sinks of degree d (Equation (4.10)), by the total number of labelled plane trees (Equation (2.2)) we get,

$$A = \frac{(n-d-1)!\frac{d}{n-1}\binom{n}{d+1}\binom{2n-d-3}{n-2}}{\frac{n!}{n}\binom{2n-2}{n-1}},$$

as the average number of root sinks of degree d in a random plane tree on n vertices. Simplifying the average we get,

$$\begin{aligned}
A = & \frac{\frac{n!(n-d-1)!d(2n-d-3)!}{(d+1)!(n-d-1)!(n-1)(n-2)!(n-d-1)!}}{\frac{(n-1)!(2n-2)!}{(n-1)!(n-1)!}} \\
= & \frac{n!d(2n-d-3)!(n-1)!}{(d+1)!(n-1)(n-2)!(n-d-1)!(2n-2)!} \\
= & \frac{n!d(2n-d-3)!}{(d+1)!(n-d-1)!(2n-2)!} \\
= & \frac{d}{(d+1)!}\frac{n(n-1)(n-2)\cdots(n-d)}{(2n-2)(2n-3)\cdots(2n-d-2)} \\
= & \frac{d}{2^{d+1}(d+1)!}\frac{n^{d+1}+\cdots}{n^{d+1}+\cdots}.
\end{aligned}$$

Taking limits as n tends to infinity we get,

$$\begin{aligned}
\lim_{n\to\infty} A &= \lim_{n\to\infty} \frac{d}{2^{d+1}(d+1)!}\frac{n^{d+1}+\cdots}{n^{d+1}+\cdots} \\
&= \frac{d}{2^{d+1}(d+1)!}.
\end{aligned}$$

□

Setting $d = 0$ in Equation (4.11) we get that the average number of root sinks of degree 0 is zero. This implies that there is no leaf sink which is also a root. For the remainder of this section, we enumerate the trees by leaf sinks.

Proposition 4.3.7. *The total number of trees of order n in which vertex j is a leaf sink reachable from root i in ℓ steps is*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell-1}. \tag{4.12}$$

Proof. The result follows by setting $d = 0$ in Proposition 4.3.1. However, to show the decomposition we will construct the proof. Let $P(x)$ to be the generating function for plane trees where x is marking a non-root vertex. Consider a plane tree rooted at vertex i such that there is a path of length ℓ starting at vertex i and terminating at a vertex of label $i+\ell$ which is also a leaf sink. The path decomposes the tree into left and right plane subtrees upto length ℓ hence we have $(P(x)xP(x))^{\ell}$ (See Figure 4.2).

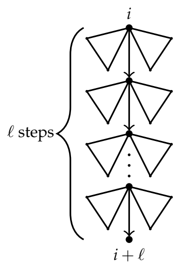


Figure 4.2: Unlabelled ordered tree with path length ℓ with vertex $i+\ell$ as a leaf sink.

Vertex j is not connected to any other tree thus it is represented by x. Putting everything together we get $x(P(x)xP(x))^{\ell} = x(xP(x)^2)^{\ell}$ as the generating function of the unlabelled trees rooted at vertex i with a path of length ℓ starting at the root and ending at leaf sink $i+\ell$. The generating function for the number of unlabelled plane trees satisfies $P(x) = \frac{1}{1-xP(x)}$. We set $xP(x) = F(x)$ so that $F(x) = \frac{x}{1-F(x)}$.

By Lagrange Inversion Formula, we obtain

$$[x^n]x(xP(x)^2)^{\ell} = [x^{n+\ell-1}]F(x)^{2\ell} = \frac{2\ell}{n+\ell-1}[t^{n-\ell-1}]\left((1-t)^{-1(n+\ell-1)}\right),$$

which is equivalent to

$$\frac{2\ell}{n+\ell-1}[t^{n-\ell-1}]\sum_{i\geq 0}\binom{-(n+\ell-1)}{i}(-t)^i$$

by Binomial Theorem. Now, applying Equation (4.2), we get

$$\begin{aligned}\frac{2\ell}{n+\ell-1}[t^{n-\ell-1}]\sum_{i\geq 0}\binom{-(n+\ell-1)}{i}(-t)^i &= \frac{2\ell}{n+\ell-1}[t^{n-\ell-1}]\sum_{i\geq 0}\binom{n+\ell+i-2}{i}t^i\\ &= \frac{2\ell}{n+\ell-1}\binom{2n-3}{n-\ell-1}\\ &= \frac{2\ell}{n+\ell-1}\binom{2n-3}{n+\ell-2}.\end{aligned}$$

Multiplying by $\frac{n-1}{n-1}$ we get

$$\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1},$$

as the formula for the number of unlabelled plane trees with a path of length ℓ starting at a root and ending at a leaf sink. Consider a path of length ℓ starting at vertex i and ending at vertex j. There are $\binom{j-i-1}{\ell-1}$ possible choices for labelling the vertices on the paths. Once the $\ell+1$ vertices on the path have been labelled, there are $(n-\ell-1)!$ ways of labelling the remaining vertices. Therefore, the total number of labelled plane trees of order n in which vertex j is a leaf sink reachable from vertex i in ℓ steps is

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell-1}.$$

This completes the proof. □

Corollary 4.3.8. *The total number of leaf sinks that are reachable from root i in ℓ steps in trees with n vertices is given by*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n-i}{\ell}\binom{2n-2}{n+\ell-1}. \tag{4.13}$$

Proof. We obtain the result by summing over all j in Equation (4.12).

$$(n-\ell-1)!\frac{\ell}{n-1}\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell-1}.$$

Let $k = j - i - 1$, so that

$$(n-\ell-1)!\frac{\ell}{n-1}\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell-1}$$
$$= (n-\ell-1)!\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1}\sum_{k=\ell-1}^{n-i-1}\binom{k}{\ell-1}.$$

By identity (4.1), we get

$$(n-\ell-1)!\frac{\ell}{n-1}\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell-1}$$
$$= (n-\ell-1)!\frac{\ell}{n-1}\binom{n-i}{\ell}\binom{2n-2}{n+\ell-1}.$$

Thus the formula. □

Corollary 4.3.9. *There are*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell-1}, \tag{4.14}$$

leaf sinks at step ℓ that are reachable from the root in trees with n vertices.

Proof. The result is evident by summing over all i in Equation (4.13). That is, the desired formula is

$$(n-\ell-1)!\frac{\ell}{n-1}\sum_{i=1}^{n-\ell}\binom{n-i}{\ell}\binom{2n-2}{n+\ell-1},$$

which we now simplify. By setting $k = n - i$, the formula becomes

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1}\sum_{k=\ell}^{n-1}\binom{k}{\ell}.$$

Now, applying identity (4.1), we get

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell-1}.$$

This is the required formula. □

Corollary 4.3.10. *The formula for the number of leaf sinks in trees of order n that are reachable from the root is*

$$\frac{n!}{n-1}\sum_{\ell=0}^{n-1}\frac{\ell}{(\ell+1)!}\binom{2n-2}{n+\ell-1}.$$

Proof. The formula follows by summing over all ℓ in Equation (4.14).

$$\begin{aligned}
&\sum_{\ell=0}^{n-1}(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell-1} \\
&\quad=\sum_{\ell=0}^{n-1}\frac{n!(n-\ell-1)!}{(\ell+1)!(n-\ell-1)!}\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1} \\
&\quad=\sum_{\ell=0}^{n-1}\frac{n!}{(\ell+1)!}\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1} \\
&\quad=\frac{n!}{n-1}\sum_{\ell=0}^{n-1}\frac{\ell}{(\ell+1)!}\binom{2n-2}{n+\ell-1}.
\end{aligned}$$

□

Corollary 4.3.11. *The average number of leaf sinks that are reachable from the root in ℓ steps in a random tree is*

$$\frac{\ell}{(\ell+1)!}.$$

Proof. Dividing the total number of leaf sinks that are reachable from the root in a labelled plane tree (See Equation (4.14)), by the total number of labelled plane trees given by Equation (2.2) we obtain,

$$B=\frac{\binom{n}{\ell+1}(n-\ell-1)!\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1}}{\frac{n!}{n}\binom{2n-2}{n-1}}$$

as the average number of leaf sinks that are reachable from the root in ℓ steps in trees on n vertices. Simple manipulations gives

$$\begin{aligned}
B=&\quad\frac{\ell}{(\ell+1)!}\frac{n!(n-2)!}{(n-\ell-1)!(n+\ell-1)!} \\
=&\quad\frac{\ell}{(\ell+1)!}\frac{(n-2)(n-3)\cdots(n-\ell)}{(n+\ell-1)(n+\ell-2)\cdots(n+\ell-(\ell-1))} \\
=&\quad\frac{\ell}{(\ell+1)!}\frac{n^{\ell-1}+\cdots}{n^{\ell-1}+\cdots}.
\end{aligned}$$

Moreover, taking limits as n tends to infinity we get

$$\begin{aligned}
\lim_{n\to\infty}B&=\lim_{n\to\infty}\frac{\ell}{(\ell+1)!}\frac{n^{\ell-1}+\cdots}{n^{\ell-1}+\cdots} \\
&=\frac{\ell}{(\ell+1)!}.
\end{aligned}$$

This is the desired result. $\square$

4.4 Enumeration by left most paths and first children

In this section, we continue our investigation of reachable vertices according to lengths of left most paths and first children. We begin by left most paths. Recall a *left most path* refers to a sequence of edges joining the eldest children at each level in a plane tree.

Proposition 4.4.1. *The number of trees of order n in which there is a left most path of length ℓ from a vertex i to a vertex j is given by*

$$(n-\ell-1)!\frac{\ell+1}{n}\binom{j-i-1}{\ell-1}\binom{2n-\ell-2}{n-\ell-1}. \qquad (4.15)$$

Proof. Let $P(x)$ be the generating function for number of plane tree. Here again, x marks the number of unrooted tree. Figure 4.3 gives the decomposition of these trees by left most path. The decomposition shows that $(xP(x))^{\ell+1} = x^{\ell+1}P(x)^{\ell+1}$ is the generating function for the number of unlabelled trees in which there is a left most path of length ℓ.

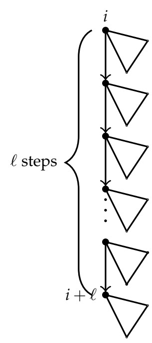


Figure 4.3: Unlabelled ordered tree with left most path of length l

It now remain to extract the coefficient of x^n in $x^{\ell+1}P(x)^{\ell+1}$. We set $xP(x) = F(x)$

so that $F(x) = \frac{x}{1-F(x)}$ and by Lagrange Inversion Formula we get

$$\begin{aligned}[x^n](xP(x))^{\ell+1} &= [x^n]F(x)^{\ell+1} \\ &= \frac{\ell+1}{n}[t^{n-\ell-1}](1-t)^{-n}.\end{aligned}$$

By Binomial Theorem and identity (4.2), we get

$$\begin{aligned}[x^n]x^{\ell+1}P(x)^{\ell+1} &= \frac{\ell+1}{n}[t^{n-\ell-1}]\sum_{k\geq 0}\binom{-n}{k}(-t)^k \quad = \frac{\ell+1}{n}[t^{n-\ell-1}]\sum_{k\geq 0}\binom{n-1+k}{k}t^k \\ &= \frac{\ell+1}{n}\binom{2n-\ell-2}{n-\ell-1}.\end{aligned}$$

There are $\binom{j-i-1}{\ell-1}$ choices for labels of the vertices on the paths of length ℓ between vertex i and vertex $i+\ell$. After the $\ell+1$ vertices on the path that have been labelled, by choice of paths, the remaining vertices are labelled in $(n-\ell-1)!$ ways. Therefore, we find that the number of trees of order n in which there is a left most path of length ℓ is given by

$$(n-\ell-1)!\frac{\ell+1}{n}\binom{j-i-1}{\ell-1}\binom{2n-\ell-2}{n-\ell-1}.$$

□

Corollary 4.4.2. *The number of trees on n vertices in which there is a left most path of length ℓ from root i is given by*

$$(n-\ell-1)!\frac{\ell+1}{n}\binom{n-i}{\ell}\binom{2n-\ell-2}{n-\ell-1}. \tag{4.16}$$

Proof. The result follows by summing over all j in Equation (4.15).

□

Corollary 4.4.3. *The formula for the number of trees of order n in which there is a left most path of length ℓ from the root is*

$$(n-\ell-1)!\frac{\ell+1}{n}\binom{n}{\ell+1}\binom{2n-\ell-2}{n-\ell-1}. \tag{4.17}$$

Proof. We obtain the formula by summing over all i in Equation (4.16).

$\square$

Moreover, summing over all ℓ, in Equation (4.17) we obtain

$$(n-1)!\sum_{\ell=0}^{n-1}\frac{1}{\ell!}\binom{2n-\ell-2}{n-\ell-1},$$

as the formula for the number of trees of order n such that there is a left most path starting at the root.

Setting $\ell = 0$ in Equation (4.17) we rediscover the formula for the number of labelled plane trees, that is $n!C_{n-1}$ where C_n is the n^{th} Catalan number.

Corollary 4.4.4. *The average number of eldest children at level ℓ in a random tree is*

$$\frac{1}{\ell!2^{\ell}}.$$

Proof. If we divide the number of labelled plane trees of order n in which there is a left most path of length ℓ (See Equation (4.17)), by the total number of labelled plane trees in Equation (2.2), we obtain

$$C = \frac{(n-\ell-1)!\frac{\ell+1}{n}\binom{n}{\ell+1}\binom{2n-\ell-2}{n-\ell-1}}{n!\frac{1}{n}\binom{2n-2}{n-1}},$$

as the average number of trees on n vertices in which there is a left most path of length ℓ. We simplify the average to get

$$\begin{aligned} C &= \frac{n!(n-\ell-1)!(\ell+1)(2n-\ell-2)!(n-1)!}{(\ell+1)!(n-\ell-1)!n(n-\ell-1)!(n-1)!(2n-2)!} \\ &= \frac{(n-1)!(2n-\ell-2)!}{\ell!(n-\ell-1)!(2n-2)!}. \end{aligned}$$

Now, taking limits as n goes to infinity, we get

$$\begin{aligned}
\lim_{n\to\infty} C &= \frac{1}{\ell!}\lim_{n\to\infty}\frac{(n-1)!(2n-\ell-2)!}{(n-\ell-1)!(2n-2)!}\\
&= \frac{1}{\ell!}\lim_{n\to\infty}\frac{(n-1)(n-2)(n-3)\cdots(n-\ell)!}{(2n-2)(2n-3)\cdots(2n-\ell-1)}\\
&= \frac{1}{\ell!}\lim_{n\to\infty}\frac{n^\ell+\cdots}{(2n)^\ell+\cdots}\\
&= \frac{1}{\ell!2^\ell}\lim_{n\to\infty}\frac{n^\ell+\cdots}{n^\ell+\cdots}\\
&= \frac{1}{\ell!2^\ell}.
\end{aligned}$$

This completes the proof. □

We switch our discussion to leaf sinks and left most paths.

Proposition 4.4.5. *The number of trees of order n rooted at vertex i in which there is a left most path of length ℓ such that the final vertex j is a leaf sink is given by*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-\ell-3}{n-\ell-1}. \tag{4.18}$$

Proof. Let $P(x)$ be the generating function for plane trees where x is marking a non-root vertex. Consider a plane tree rooted at vertex i such that there is a left most path of length ℓ starting at vertex i and ending at vertex $i+\ell$ which is also a leaf sink. The path decomposes the tree into right plane subtrees upto length ℓ hence we have $(xP(x))^\ell$ (See Figure 4.4).

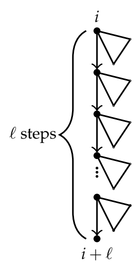


Figure 4.4: Unlabelled ordered tree with left most path of length ℓ and the final vertex is a leaf sink.

Vertex $i+\ell$ is not connected to any other tree thus it is represented by x. Putting everything together, we get $x(xP(x))^{\ell} = x^{\ell+1}P(x)^{\ell}$ as the generating function of the unlabelled plane tree rooted at vertex i with a left most path of length ℓ starting at the root and ending at leaf sink $i+\ell$. We set $xP(x) = F(x)$ and applying Lagrange Inversion Formula, to get

$$\begin{aligned}
[x^n]x^{\ell+1}P(x)^{\ell} &= [x^n]xF(x)^{\ell} \\
&= [x^{n-1}]F(x)^{\ell} \\
&= \frac{\ell}{n-1}[t^{n-\ell-1}](1-t)^{-(n-1)}.
\end{aligned}$$

This is equivalent to

$$\frac{\ell}{n-1}[t^{n-\ell-1}]\sum_{k\geq 0}\binom{-(n-1)}{k}(-t)^k$$

by Binomial Theorem 4.1.1. Now, by Equation (4.2), and substituting for $k = n-\ell-1$, we get

$$\frac{\ell}{n-1}\binom{2n-\ell-3}{n-\ell-1},$$

as the formula for the number of unlabelled plane trees with a left most path of length ℓ starting at a root and ending at a leaf sink. Consider a path of length ℓ starting at vertex i and ending at vertex j. There are $\binom{j-i-1}{\ell-1}$ possible choices for labelling the vertices on the paths. Once the $\ell+1$ vertices on the path have been labelled, there are $(n-\ell-1)!$ ways of labelling the remaining vertices. Therefore, putting everything together we obtain the desired formula. □

Corollary 4.4.6. *The number of trees on n vertices in which there is a left most path of length ℓ from root i and a final vertex is a leaf sink is given by*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n-i}{\ell}\binom{2n-\ell-3}{n-\ell-1}. \tag{4.19}$$

Proof. We get the result by summing over all j in equation (4.18). □

By summing over all i in Equation (4.19) we get

Corollary 4.4.7. *The total number of trees on n vertices in which there is a left most path of length ℓ from the root and the final vertex is a leaf sink is given by*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-\ell-3}{n-\ell-1}. \tag{4.20}$$

Corollary 4.4.8. *The number of trees of order n in which there is a left most path starting from the root and the ending vertex is a leaf sink is given by*

$$n(n-2)!\sum_{\ell=0}^{n-1}\frac{\ell}{(\ell+1)!}\binom{2n-\ell-3}{n-\ell-1}.$$

Proof. We sum over all ℓ in Equation (4.20) and then simplify to obtain the desired result. □

Corollary 4.4.9. *The average number of eldest children which are also leaf sinks in random tree is given by*

$$\frac{1}{(\ell+1)!2^{\ell+1}}.$$

Proof. We divide the total number of labelled plane trees in which there is a left most path of length ℓ from the root and the final vertex is a leaf sink (See Equation (4.20)), by the total number of labelled plane trees given by Equation (2.2) to obtain

$$\frac{(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-\ell-3}{n-\ell-1}}{n!\frac{1}{n}\binom{2n-2}{n-1}}$$

as the average number of labelled plane trees in which there is a left most path and a final vertex is a leaf sink. We simplify and tend $n \to \infty$ to obtain the desired result. □

Recall that in ordered trees, the children (or sibling) are linearly ordered and they are drawn in a left -right pattern where the left most child is called the *first child* to the parent. Therefore, enumerating the trees by first children we find that,

Proposition 4.4.10. *The number of trees of order n with a vertex i as a root and vertex j as a first child reachable from root i in ℓ steps is given by*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell-1}.$$

Proof. Let $P(x)$ be the generating function for the plane trees where x represents non-root vertices. Consider a plane tree rooted at vertex i such that there is a path of length ℓ starting at vertex i and terminating at vertex $i+\ell$ which is a first child. The path decomposes the tree into left and right plane subtrees upto vertex $\ell-1$ hence we have $(xP(x)^2)^{\ell-1}$. (See Figure 4.5). Since vertex $i+\ell$, which is the $(\ell+1)^{th}$ vertex, is a first child, it's parent has no left subtree. Vertex $i+\ell$ can either have children or not. Thus the decomposition gives $(x(P(x)^2)^{\ell-1}xP(x)xP(x) = x^{\ell+1}P(x)^{2\ell}$ as the generating function for unlabelled plane trees rooted at vertex i with a path of length ℓ starting at the root i and ending at a first child $i+\ell$. This is pictorially represented as

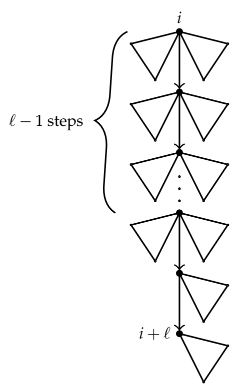


Figure 4.5: Unlabelled ordered tree of order n with first child at step ℓ.

By Lagrange Inversion Formula, we get

$$\begin{aligned}[x^n]\left(x^{\ell+1}P(x)^{2\ell}\right) &= [x^{n+\ell-1}]F(x)^{2\ell}\\ &= \frac{2\ell}{n+\ell-1}[t^{n-\ell-1}](1-t)^{-(n+\ell-1)},\end{aligned}$$

which simplifies to

$$\frac{2\ell}{n+\ell-1}[t^{n-\ell-1}]\sum_{i\geq 0}\binom{-(n+\ell+i-1)}{i}-(t)^i$$

using Binomial Theorem. Now, by identity (4.2), we obtain

$$\begin{aligned}
\frac{2\ell}{n+\ell-1}[t^{n-\ell-1}]\sum_{i\geq 0}\binom{-(n+\ell+i-1)}{i}-(t)^i &= \frac{2\ell}{n+\ell-1}[t^{n-\ell-1}]\sum_{i\geq 0}\binom{n+\ell+i-2}{i}t^i \\
&= \frac{2\ell}{n+\ell-1}\binom{2n-3}{n-\ell-1} \\
&= \frac{2\ell}{n+\ell-1}\binom{2n-3}{n+\ell-2}.
\end{aligned}$$

Multiplying by $\frac{n-1}{n-1}$, we get

$$\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1}$$

as the formula for unlabelled plane trees with a path of length ℓ starting at a root and terminating at a first child. The number of choices for labelling the vertices on a path of length ℓ between vertex i and vertex j is $\binom{j-i-1}{\ell-1}$. Since the $\ell+1$ vertices along the path have been labelled, then the remaining vertices can be labelled in $(n-\ell-1)!$ ways. Finally, by assembling everything together we complete the proof by obtaining the required formula as

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell-1}.$$

□

Corollary 4.4.11. *The number of first children at level ℓ that are reachable from root i in trees on n vertices is given by*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n-i}{\ell}\binom{2n-2}{n+\ell-1}. \tag{4.21}$$

Proof. We obtain the required result by summing over all j in Equation (4.20). □

Corollary 4.4.12. *The number of first children at level ℓ in trees of order n that are reachable from the root is given by*

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell-1}. \tag{4.22}$$

Proof. Summing over all i in Equation (4.21), we obtain

$$\begin{aligned}(n-\ell-1)!\frac{\ell}{n-1}\sum_{i=1}^{n-l}\binom{n-i}{\ell}\binom{2n-2}{n+\ell-1}\\ = (n-\ell-1)!\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1}\sum_{i=1}^{n-\ell}\binom{n-i}{\ell},\end{aligned}$$

which simplifies to

$$(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell-1}$$

upon use of identity (4.1). □

Corollary 4.4.13. *The total number of first children that are reachable from the root in trees of order n is given by*

$$\frac{n!}{n-1}\sum_{\ell=0}^{n-1}\frac{\ell}{n-1}\binom{2n-2}{n+\ell-1}.$$

Proof. We get the result by summing over all ℓ in Equation (4.22) then simplifying. □

Corollary 4.4.14. *The average number of first children that are reachable from the root in ℓ steps in a random tree is*

$$\frac{\ell}{(\ell+1)!}.$$

Proof. By dividing the total number of first children in a labelled plane tree that are reachable from the root (See Equation (4.22)), by the total number of labelled plane trees given by Equation (2.2) we obtain,

$$\frac{(n-\ell-1)!\frac{\ell}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell-1}}{\frac{n!}{n}\binom{2n-2}{n-1}}$$

as the average number of first children in a labelled plane tree of order n that are reachable from the root in ℓ steps. We tend $n \to \infty$ to obtain the required result. □

Remark 4.4.15. The number of unlabelled plane trees with a path of length ℓ starting at a root i and terminating at a leaf sink j has a similar generating function as for the number of unlabelled trees with a path of length ℓ starting at a root i and terminating at a first child j. Therefore, they pose similar results if we sum over i, j and ℓ. Asymptotic results are also the same.

Remark 4.4.16. If terminating vertex is a first child which is also a leaf, we obtain generating function as $(xP(x)^2)^{\ell-1}xP(x)x$ which gives

$$\frac{2\ell-1}{n+\ell-1}\binom{2n-3}{n+\ell-2}$$

as the number of such trees on n vertices.

4.5 Enumeration by non-first children and non-leaf sinks

In plane trees, any vertex which is not an eldest child is called a *non first child*. A vertex which is not a leaf sink is a non leaf sink. In this section, we enumerate trees with respect to number of non first children as well as number of non leaf sinks.

Proposition 4.5.1. *The number of trees on n vertices rooted at vertex i and having vertex j as a non first child which is reachable in ℓ steps is given by*

$$(n-\ell-1)!\frac{\ell+1}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell}. \tag{4.23}$$

Proof. We construct the generating function by considering a plane tree rooted at vertex i with a path of length ℓ starting at i and terminating at $i+\ell$, which is a non first child. The path decomposes the tree into left and right plane subtrees as shown in Figure 4.6.

The decomposition gives $x^{\ell+2}P(x)^{2\ell+2}$ as the generating function for unlabelled plane trees, with a path of length ℓ starting at a root and terminating at a non

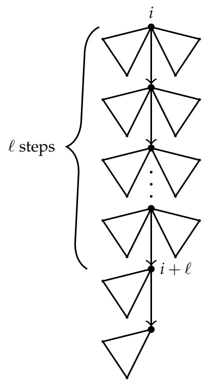


Figure 4.6: Unlabelled ordered tree with non first children at level ℓ.

first child. By Lagrange Inversion formula, we have

$$\begin{aligned}[x^n]x^{\ell+2}P(x)^{2\ell+2} &= [x^{n+\ell}]F(x)^{2\ell+2}\\ &= \frac{2\ell+2}{n+\ell}[t^{n-\ell-2}]\left((1-t)^{-1(n+\ell)}\right).\end{aligned}$$

Applying Binomial Theorem 4.1.1 and identity (4.2) we get

$$\begin{aligned}\frac{2\ell+2}{n+\ell}[t^{n-\ell-2}]\left((1-t)^{-1(n+\ell)}\right) &= \frac{2\ell+2}{n+\ell}[t^{n-\ell-2}]\sum_{i\geq 0}\binom{n+\ell+i-1}{i}t^i\\ &= \frac{2\ell+2}{n+\ell}\binom{2n-3}{n-\ell-2}\\ &= \frac{2\ell+2}{n+\ell}\binom{2n-3}{n+\ell-1}.\end{aligned}$$

Hence multiplying by $\frac{n-1}{n-1}$ we get

$$\frac{\ell+1}{n-1}\binom{2n-2}{n+\ell},$$

as the formula for the number of non first children at step ℓ that are reachable from the root in trees of order n. The choices for labels of the vertices on a path of length ℓ from vertex i to vertex j is $\binom{j-i-1}{\ell-1}$. After the $\ell+1$ vertices on the path have be labelled, there are $(n-\ell-1)!$ possible ways of labelling the remaining vertices. Therefore, the total number of trees of order n in which vertex j is a non first child reachable from i in ℓ steps is

$$(n-\ell-1)!\frac{\ell+1}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell}.$$

This completes the proof. $\square$

Corollary 4.5.2. *The total number of non first children at level ℓ in trees of order n that are reachable form vertex i is*

$$(n-\ell-1)!\frac{\ell+1}{n-1}\binom{n-i}{\ell}\binom{2n-2}{n+\ell}. \tag{4.24}$$

Proof. The required formula is obtained by summing over all j in Equation (4.23). $\square$

Corollary 4.5.3. *The total number of non first children at level ℓ that are reachable from the root in a tree on n vertices is given by*

$$(n-\ell-1)!\frac{\ell+1}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell}. \tag{4.25}$$

Proof. Summing over all i in Equation (4.24), we obtain the said formula. $\square$

Corollary 4.5.4. *The total number of non first children in trees of order n that are reachable from the root is given by*

$$\frac{n!}{n-1}\sum_{\ell=0}^{n-1}\frac{1}{\ell!}\binom{2n-2}{n+\ell}.$$

Proof. We sum over all ℓ in Equation (4.25), and simplify to arrive at the desired formula. $\square$

Corollary 4.5.5. *The average number of non first children that are reachable from the root in ℓ steps in a random tree is given by $\frac{1}{\ell!}$.*

Proof. We divide the total number of non first children at level ℓ that are reachable from the root in a labelled plane tree (See equation (4.25)) by the total number of labelled plane trees given by (2.2) to get

$$\frac{(n-\ell-1)!\frac{\ell+1}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell}}{n!\frac{1}{n}\binom{2n-2}{n-1}}.$$

This is the average number of non first children at level ℓ that are reachable from the root in trees on n vertices. On further simplification and tending $n \to \infty$ we get the required formula as $\frac{1}{\ell!}$. $\square$

For the remainder of this section, we enumerate non-leaf sinks.

Proposition 4.5.6. *The number of trees of order n with vertex i as a root and vertex j as non leaf sink reachable from the root in ℓ steps is given by*

$$(n-\ell-1)!\frac{\ell+1}{n-1}\binom{j-i-1}{\ell-1}\binom{2n-2}{n+\ell}. \tag{4.26}$$

Proof. Consider a plane tree rooted at vertex i with a path of length ℓ to vertex $i+\ell$ (See Figure 4.7).

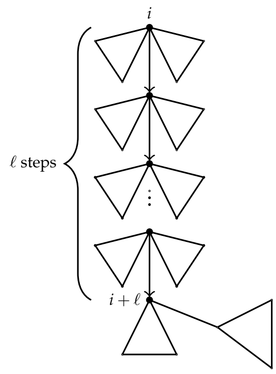


Figure 4.7: Unlabelled ordered tree with non leaf sink at level ℓ.

The path divides the tree into left and right plane subtrees upto ℓ steps, $(xP(x)^2)^\ell$. Since vertex $i+\ell$ is a non leaf sink, it must therefore have an outdegree greater than zero. That is, it can be a parent to a plane subtree which can possibly be empty hence we represent it as $xP(x)$ or it can have another sibling which is having other plane subtrees represented by $xP(x)$. Therefore putting everything together we obtain $\left(xP(x)^2\right)^\ell xP(x)xP(x) = x^{\ell+2}P(x)^{2\ell+2}$ as the generating function for the number of trees with a path of length ℓ starting from a root and ending at a non leaf sink. As before we set $xP(x) = F(x)$ and apply Lagrange

Inversion Formula to get

$$[x^n]x^{\ell+2}P(x)^{2\ell+2} = [x^{n+\ell}]F(x)^{2\ell+2} \qquad = \frac{2\ell+2}{n+\ell}[t^{n-\ell-2}]\left((1-t)^{-1(n+\ell)}\right).$$

Using Binomial theorem 4.1.1 and Equation (4.2), we obtain

$$\begin{aligned}
\frac{2\ell+2}{n+\ell}[t^{n-\ell-2}]\left((1-t)^{-1(n+\ell)}\right) &= \frac{2\ell+2}{n+\ell}[t^{n-\ell-2}]\sum_{i\geq 0}\binom{n+\ell+i-1}{i}t^i \\
&= \frac{2\ell+2}{n+\ell}\binom{2n-3}{n-\ell-2} \\
&= \frac{2\ell+2}{n+\ell}\binom{2n-3}{n+\ell-1}.
\end{aligned}$$

Multiplying by $\frac{n-1}{n-1}$ we get

$$\frac{\ell+1}{n-1}\binom{2n-2}{n+\ell}.$$

This is the formula for the number of unlabelled trees on n vertices with a path of length ℓ starting at the root and ending at non leaf sink. □

Corollary 4.5.7. *The total number of non leaf sinks at level ℓ in trees of order n that are reachable from root i is*

$$(n-\ell-1)!\frac{\ell+1}{n-1}\binom{n-i}{\ell}\binom{2n-2}{n+\ell}. \tag{4.27}$$

Proof. The result follows immediately by summing over all j in equation (4.26). □

Corollary 4.5.8. *The total number of non leaf sinks at level ℓ that are reachable from the root in trees with n vertices is*

$$(n-\ell-1)!\frac{\ell+1}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell}. \tag{4.28}$$

Proof. We sum over all i in Equation (4.27), to obtain the result. □

Corollary 4.5.9. *The total number of non leaf sinks that are reachable in trees of order n is*

$$\frac{n!}{n-1}\sum_{\ell=0}^{n-1}\frac{1}{\ell!}\binom{2n-2}{n+\ell}.$$

Proof. We obtain the formula by summing over all ℓ in equation (4.28). $\square$

Corollary 4.5.10. *There are on average $\frac{1}{\ell!}$ non leaf sinks that are reachable in ℓ steps in random trees.*

Proof. If we divide the total number of non leaf sinks at level ℓ that are reachable from the root in a labelled plane tree of order n (See Equation (4.28)), by the total number of labelled plane trees in Equation (2.2) we get

$$\frac{(n-\ell-1)!\frac{\ell+1}{n-1}\binom{n}{\ell+1}\binom{2n-2}{n+\ell}}{n!\frac{1}{n}\binom{2n-2}{n-1}}$$

as the average number of non leaf sinks at level ℓ that are reachable from the root. By simple manipulation and taking $n \to \infty$, we obtain the required result. $\square$

Remark 4.5.11. The number of unlabelled plane trees with a path of length ℓ starting at a root i and terminating at a non leaf sink j has similar generating function as the number of unlabelled trees with a path of length ℓ starting at a root i and terminating at a non first child j. Therefore they pose similar results if we sum over i, j and ℓ. Asymptotic results are the same.

4.6 Enumeration by exact number of vertices

The number of trees in which an exact number of vertices are reachable from a given vertex have been studied for the case of labelled ordinary trees as well as non crossing trees. Quite a number of results were obtained by Okoth in his PhD thesis, [14]. Similarly, Seo and Shin [20] established a formula for rooted Cayley trees in which a maximal increasing subtree of order k has exactly k reachable vertices. In this section we investigate the exact number of vertices that are reachable from vertex 1 in labelled ordered trees.

Theorem 4.6.1. *The total number of trees of order n such that exactly k vertices are reachable from the root is given by*

$$O_{n,k} = \sum_{k \le m+1 \le n} \binom{n}{m+1} z(m, k-1) \frac{m-k+1}{n-k} (n-k)^{(n-m-1)} \tag{4.29}$$

for $0 \le k < n$, $O(n,n) = (2n-3)!!$, *where* $n^{(r)} = n(n+1)(n+2)\cdots(n+r-1)$ *is a rising factorial and $z(m,k)$ is the number of ordered trees on $m+1$ vertices with additional $(m-k)$ decreasing leaves attached to an increasing tree with k edges. The trees counted by $z(m,k)$ were enumerated by Drake [8] .*

Proof. A subtree of a v-rooted tree T is said to be *increasing* if the labels in the subtree are increasing as the vertices move away from the root. A *maximal decreasing* subtree, is a decreasing subtree rooted at vertex v with the highest number of vertices. Seo and Shin [20], showed that Equation (4.29) gives the number of ordered trees on $[n] := \{1, 2, \ldots\}$ with its maximal decreasing ordered subtree having k vertices. Now, orienting the edges from vertices of lower label to vertices of higher label in an ordered tree, we obtain a maximal increasing subtree of order n having exactly k reachable vertices from the root. □

Corollary 4.6.2. *The number of trees of order n having exactly $k \ge 2$ vertices reachable from vertex* 1 *is given by*

$$O_{n-1,k-1}(2k-3) \tag{4.30}$$

where $O_{n,k}$ is given by Equation (4.29).

Proof. Consider an ordered tree P with $n-1$ vertices of labels $2, 3, \cdots, n$ rooted at vertex v_1 and having exactly $k-1$ vertices reachable from the root. There are $O_{n-1}, k-1$ such trees. We follow the steps below to obtain trees of order n in which k vertices are reachable from vertex 1:

Step 1: Let P_0 be the maximal increasing subtree having $v_1, v_2, \ldots, v_{k-1}$ vertices where $v_i < v_{i+1}$ for all i. In P, deleting all the edges in P_0 we obtain non-single vertex subtrees $P_1, P_2, \cdots, P_m$.

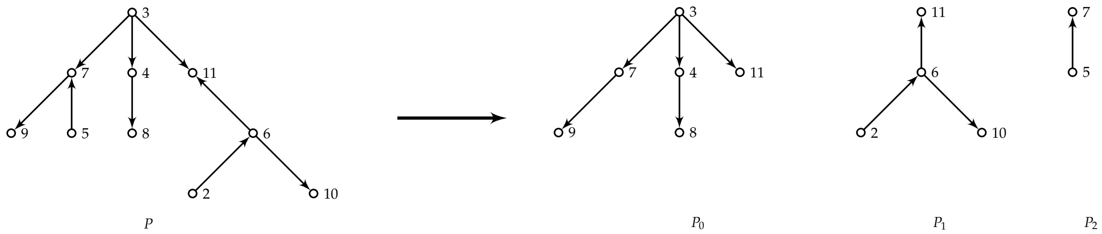


Figure 4.8: Diagram showing Step 1 in the proof of Corollary 4.6.2

Step 2: Relabelling the vertices of the maximal increasing subtree P_0 with vertex v_1 now as 1, v_2 as v_1, v_3 as v_2 and so on, the maximal increasing tree P_0 remain with $k-1$ vertices. There are $2k-3$ positions in the new maximal increasing subtree rooted at 1 where vertex v_{k-1} can be attached. For each maximal increasing subtree previously rooted at v_1, we obtain $2k-3$ new subtrees rooted at vertex 1 with k reachable vertices.

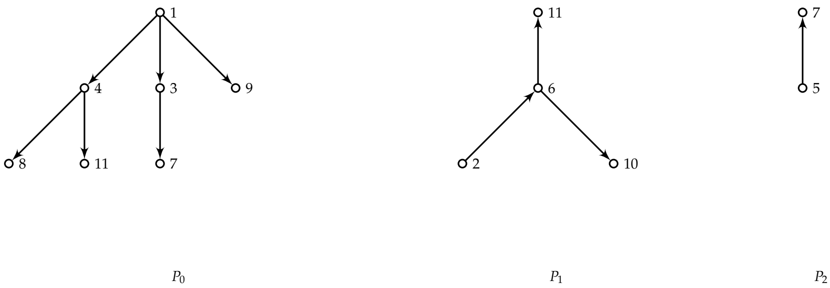


Figure 4.9: Diagram showing Step 2 in the proof of Corollary 4.6.2

Step 3: Identify vertex v_i in the subtrees $P_1, P_2, \ldots, P_m$ with vertex v_i in the new maximal increasing subtree, for all $i \in \{1, \ldots, m\}$. We obtain a tree of order n in which exactly k vertices are reachable from the root labelled 1.

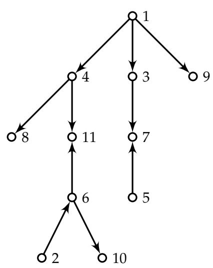


Figure 4.10: Diagram showing Step 3 in the proof of Corollary 4.6.2

Thus there are $(2k-3)O_{n-1,k-1}$ trees of order n rooted at vertex 1 and having exactly k vertices reachable from the root.

$\square$

Corollary 4.6.3. *Let $P(n,i)$ be the total number of trees on n vertices with exactly $n-i+1$ reachable vertices from vertex i. We have*

$$P(n,i) = (2n-2i+1)!!(n-i+1)(n+1)(n+2)\cdots(n+i-2).$$

Proof. There are $(2n-2i+1)!!$ recursive trees on $n-i+1$ vertices (See Lemma 2 in [20]). Since there are $n-(n-i+1) = i-1$ vertices which are not reachable from vertex i, then all the $i-1$ vertices have labels less than i. The number of ways of adding the $i-1$ vertices to recursive tree successively is given by $(n-i+1)(n+1)(n+2)\cdots(n+i-2)$, (See Lemma 2 in [20]). Therefore, the total number of trees on n vertices with exactly $n-i-1$ vertices reachable from vertex i is given by

$$(2n-2i+1)!!(n-i+1)(n+1)(n+2)\cdots(n+i-2).$$

Hence the desired formula. $\square$

4.7 Reachable vertices

Here we are interested in the number of reachable vertices from a vertex i which is not necessarily a root.

Lemma 4.7.1. *The number of trees with n vertices such that vertex j is reachable from vertex i in 1 step is given by*

$$\frac{2(n-1)!}{n}\binom{2n-2}{n-1}. \tag{4.31}$$

Proof. Consider a tree with n vertices. The tree has $n-1$ edges. From the tree we can choose an edge arbitrarily in $(n-1)$ ways. We label the end points of the edge as either i or j. This can be done in two ways. Therefore we have $2(n-1)$ possibilities for choosing an edge and labelling the end points as i and j. Here vertex j will be reachable from vertex i in one step. The remaining vertices can be labelled in $(n-2)!$ ways. Thus there are $2(n-1)(n-2)!$=$2(n-1)!$ possibilities for labelling any two adjacent vertices by i and j, and the remaining vertices by $\{1, 2, \cdots, n\} \setminus \{i, j\}$. The number of unlabelled trees with n vertices is given by $\frac{1}{n}\binom{2n-2}{n-1}$. Therefore the number of trees of order n in which a given vertex j is reachable from a given vertex i in one step is given by

$$\frac{2(n-1)!}{n}\binom{2n-2}{n-1}.$$

Thus the proof is complete. □

We also conjecture that:

Conjecture 4.7.2. The number of trees in which vertex 3 is reachable from vertex 1 in two steps is given by

$$(n-2)! \sum_{k=0}^{n-3} \frac{1}{k+1} \binom{2k}{k} \binom{k+1}{n-k-3}.$$

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

In this thesis, we have studied enumeration of trees with local orientation with respect to their indegree sequences and outdegree sequences as well as reachability of vertices in labelled ordered trees. We have obtained a recurrence relation satisfied by labelled trees with local orientation whose indegree sequences are given by $0^1 1^{n-1}$ and outdegree sequences by $0^{f_0} 1^{f_1} 2^{f_2} \cdots p^{f_p}$ (Theorem 3.1.2). In Chapter 4, we considered reachability of vertices in labelled plane trees. We have obtained formulas which count the number of vertices that are reachable from the root, with respect to; path length (Theorem 4.2.1), sinks (Proposition 4.3.1), leaf sinks (Proposition 4.3.7), left most paths (Proposition 4.4.1), first children (Proposition 4.4.10), non first children (Proposition 4.5.1), and non leaf sinks (Proposition 4.5.6). In each case, the average number of reachable vertices were obtained for any random plane tree. Moreover, a formula for the number of trees with an exact number of reachable vertices has also been obtained in Theorem 4.6.1. In Lemma 4.7.1, we obtained the number of plane trees of order n such that a given vertex j is reachable from vertex i in one step. Here vertex i is not necessarily a root. Our results have added to the rich literature in this area of research.

5.2 Recommendations

A *d-ary tree* is a rooted tree in which each vertex has no more than d children. A *k-noncrossing tree* is a non-crossing tree where each vertex has labels in the set $\{1, 2, \ldots, k\}$ such that the sum of labels for any two adjacent vertices in a path from the root does not exceed $k + 1$. A *k-plane tree* is a plane tree such that if the vertices are to be labelled with integers in the set $\{1, 2, \cdots, k\}$ then the sum of the labels between any two adjacent vertices in a path does not exceed $k + 1$. Reachability has not been studied in the setting of *k-plane trees*, *k-noncrossing trees* and *d-ary trees*. Similarly, the closed formulas for exact number of labelled plane trees of order n in which a given vertex j is reachable from a given vertex i in $\ell \geq 2$ steps have not been established. We, therefore, recommend that a similar study be conducted to obtain the number of reachable vertices in *k-plane* trees, *k-noncrossing* trees and *d-ary* trees. Also, a similar research should be conducted so as to obtain a closed formula for the number of trees of order n in which a given number of vertices are reachable from a specified root. Moreover, we recommend that research be conducted to obtain a formula for the number of vertices that are reachable from a given vertex i in plane trees such that vertex i is not necessarily a root.

REFERENCES

[1] C. W. Borchardt. Ueber eine der Interpolation entsprechende Darstellung der Eliminations-Resultante. J. reine angew. Math., 57:111–121,1860.

[2] A. Cayley. A theorem on trees. Quart. J. Math., 23:376–378, 1889.

[3] S. Chaiken. A Combinatorial proof of all minor matrix theorems. SIAM. J. Alg. Disc method 3, 319–329, 1982.

[4] W. Y. Chen. A general bijective algorithm for trees. Proc. Nati. Acad. Sci. USA, 87:9635–9639, 1990.

[5] L. Comtet. Advanced Combinatorics: The art of Finite and Infinite Expansions. D. Reidel Publishing Company, 1974.

[6] E. Cotterill. Geometry of curves with exceptional secant planes: linear series along the general curve. Math. Z., 267(3-4):549–582, 2011.

[7] R. Diestel. *Graph Theory*, vol. 173 of Graduate Texts in Mathematics, 3rd ed. Springer-Verlag, Berlin, 2005.

[8] B. Drake. An inversion for labelled trees and some limits of areas under lattice paths. ProQuest LLC, Ann Arbor, MI, 2008. Thesis (Ph.D.)-Brandeis University.

[9] M. Drmota. Random trees: an interplay between Combinatorics and Probability. Springer Science and Business, 2009.

[10] R. R. X. Du and J. Yin. Counting labelled trees with a given indegree sequence. J. Combin. Theory, Ser. A, 117(3):345–353, 2010.

[11] S-P. Eu, S. Seo and H. Shin. Enumeration of vertices among all rooted ordered trees with levels and degrees. arXiv:1605.00715v3 [math.CO].

[12] A. Joyal. Une théorie combinatoire des séries formelles. Adv. Math., 42:1-82, 1981.

[13] J. L. Martin and V. Reiner. Factorization of some weighted spanning tree enumerators. J. Combin. Theory, Ser. A, 104(2):287–300, 2003.

[14] I. O. Okoth. Combinatorics of oriented trees and tree like structures, PhD thesis, Stellenbosch University, 2015.

[15] I. O. Okoth and S. Wagner. Locally oriented noncrossing trees. Electron . J . Combin., 15, 2015.

[16] J. Pitman. Coalescent random forests. J. Combin. Theory, Ser. A, 85:165–193, 1999.

[17] A. Postnikov. Intransitive trees. J. Combin. Theory, Ser. A, 79:360–366,1997.

[18] H. Prüfer. Neuer Beweis eines Satzes über Permutationen. Archiv der Math. u. Physik, 3(47):142–144, 1918.

[19] J. B. Remmel and S. G.Williamson. Spanning trees and function classes. Electron. J. Combin., 9(1):229–233, 2002.

[20] S. Seo and H. Shin. Refinement for ordered labelled tree. arXiv:1207.3291v1 [math.CO].

[21] H. Shin and J. Zeng. A bijective enumeration of labelled trees with given indegree sequence. J. Combin. Theory, Ser. A, 118(1):115–128, 2011.

[22] C. G. Small. Functional equations and how to solve them. Springer Verlag, Berlin, 8, 2007.

[23] R. P. Stanley. *Enumerative Combinatorics*. Vol. 2, volume 62 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1999.

[24] J. J. Sylvester. On the change of systems of independent variables. Quart. J. Math., 1:126–134, 1857.

[25] S. Wagner. Labelled trees, functions, and an algebraic identity. Electron. J. Combin., 18, 2011.

[26] H. S. Wilf. *Generatingfunctionology*. A. K. Peters, Ltd., Natick, MA, USA, 2006.