SpringerWienNewYork

Michael Drmota

Random Trees

An Interplay between Combinatorics and Probability

SpringerWienNewYork

Univ.-Prof. Dipl.-Ing. Dr. techn. Michael Drmota
Institute of Discrete Mathematics and Geometry, Vienna University
of Technology, Austria

Printed in Germany

SpringerWienNewYork is part of Springer Science + Business Media
springer.at

Typesetting: Camera ready by the author
Printing: Strauss GmbH, 69509 Mörlenbach, Germany

Printed on acid-free and chlorine-free bleached paper

With 72 black & white figures

SPIN: 12161386

Library of Congress Control Number: 2008942098

ISBN 978-3-211-75355-2 SpringerWienNewYork

To Gabriela, Heidi, Hanni and Peter

Preface

Trees are a fundamental object in graph theory and combinatorics as well as a basic object for data structures and algorithms in computer science. During the last years research related to (random) trees has been constantly increasing and several asymptotic and probabilistic techniques have been developed in order to describe characteristics of interest of large trees in different settings.

The purpose of this book is to provide a thorough introduction into various aspects of trees in random settings and a systematic treatment of the involved mathematical techniques. It should serve as a reference book as well as a basis for future research. One major conceptual aspect is to connect combinatorial and probabilistic methods that range from counting techniques (generating functions, bijections) over asymptotic methods (singularity analysis, saddle point techniques) to various sophisticated techniques in asymptotic probability (convergence of stochastic processes, martingales). However, the reading of the book requires just basic knowledge in combinatorics, complex analysis, functional analysis and probability theory of master degree level. It is also part of concept of the book to provide full proofs of the major results even if they are technically involved and lengthy.

Due to the diversity of the topic of the book it is impossible to present an exhaustive treatment of all known models of random trees and of all important aspects that have been considered so far. For example, we do not deal with the simulation of random trees. The choice of the topics reflects the author's taste and experience. It is slightly leaning on the combinatorial side and analytic methods based on generating functions play a dominant role in most of the parts of the book. Nevertheless, the general goal is to describe the limiting behaviour of large trees in terms of continuous random objects. This ranges from central (or other) limit theorems for simple tree statistics to functional limit theorems for the shape of trees, for example, encoded by the horizontal or vertical profile. The majority of the results that we present in this book is very recent.

There are several excellent books and survey articles dealing with some aspects on combinatorics on trees and graphs resp. with probabilistic methods in these topics which complement the present book. One of the first ones was Harary and Palmer book *Graphical enumeration* [98]. Around the same time Knuth published the first three volumes of *The Art of Computer Programming* [128, 129, 130] where several classes of trees related to algorithms from computer science are systematically investigated. His books with Green *Mathematics for the analysis of algorithms* [96] and the one with Graham and Patashnik *Concrete Mathematics* [95] complement this programme. In parallel asymptotic methods in combinatorics, many of them based on generating functions, became more and more important. The articles by Bender *Asymptotic methods in enumeration* [7] and Odlyzko *Asymptotic enumeration methods* [165] are excellent surveys on this topic. This development is highlighted by Flajolet and Sedgewick's recent (monumental) monograph *Analytic Combinatorics* [84]. Computer science and in particular the mathematical analysis of algorithms was always a driving force for developing concepts for the asymptotic analysis of trees (see also the books by Kemp [122], Hofri [102], Sedgewick and Flajolet [191], and by Szpankowski [197]). Moreover, several concepts of random trees arose naturally in this scientific process (see for example Mahmoud's book *Evolution of random search trees* [146], and Pittel's, Devroye's or Janson's work).

However, combinatorics and problems of computer science, though important, are not the only origin of random tree concepts. There was at least a second (and almost independent) line of research concerning conditioned Galton-Watson trees. Here one starts with a Galton-Watson branching process and conditions on the size of the resulting trees. For example, Kolchin's book *Random Mappings* [132] summarises many results from the *Russian school*. This work is complemented by the *American school* represented by Aldous [3, 5] and Pitman [171] where stochastic processes related to the Brownian motion play an important role. The invention of the continuum random tree as well as the ISE (integrated super-Brownian excursion) by Aldous are breakthroughs. Actually these continuous limit objects are quite universal concepts. It seems that they also appear as limit objects for several kinds of random planar maps and other related discrete objects. There are even more general settings where Lévy processes are used (see the recent survey articles *Random Trees and Applications* [135] and *Random Real Trees* [136] by Le Gall and the book *Probability and Real Trees* [75] by Evans). By the way, the study of random graphs is completely different from that of random trees (compare with the books by Bollobás [21], Janson, Łuczak and Ruciński [116], and Kolchin [133]). Nevertheless, there is a very interesting paper *The Birth of the Giant Component* [115] which uses analytic methods that are very close to tree methods.

This book is divided into nine chapters. The first two of them are providing some background whereas the remaining chapters 3–9 are devoted to more specific and (more or less) self contained topics on random trees and on related subjects. Of course, they will use basic notions from Chapter 1 and some of the methods from Chapter 2.

In Chapter 1 we survey several classes of random trees that are considered here: combinatorial tree classes like planted plane trees, Galton-Watson trees, recursive trees, and search trees including binary search trees and digital trees.

Chapter 2 is a second introductory chapter. It collects some basic facts on combinatorics with generating functions and provides an analytic treatment of generating functions that satisfy a functional equation (or a system of functional equations) leading to asymptotics and central limit theorems. It is probably not necessary to study all parts of this chapter in a first reading but to use it as a reference chapter.

The first purpose of Chapter 3 is tree counting, to obtain explicit formulas for the numbers of trees of given size with possible and asymptotic information on these numbers in those cases, where no or no simple explicit formula is available. The analysis of several combinatorial classes of trees and also of Galton-Watson trees is based on generating functions and their analytic properties that are discussed in Chapter 2. The recursive structure of (rooted) trees usually leads to a functional equation for the corresponding generating functions. By extending these counting procedures with the help of bivariate generating functions one can also study (so-called) additive statistics on these tree classes like the number of nodes of given degree or more generally the number of occurrences of a given pattern. In all these cases we derive a central limit theorem.

The general topic of Chapters 4–7 is the limiting behaviour of the profile and related statistics of different classes of random trees. Starting from a natural (vertex) labelling on a discrete object, for example the distance to a root vertex in a tree, the profile is the value distribution of the labels. More precisely, if a random discrete object has size n then the profile $(X_{n,k})$ is given by the numbers $X_{n,k}$ of vertices with label k. The idea behind is that the profile $(X_{n,k})$ describes the shape of the random object. It is therefore natural to search for a proper limiting object of the profile after a proper scaling.

In Chapter 4 we discuss the depth profile (induced by the distance to the root) of Galton-Watson trees with bounded offspring variance which can be approximated by the local time of the Brownian excursion of duration 1. This property is closely related to the convergence of normalised Galton-Watson trees to the continuum random tree introduced by Aldous [2, 3, 4]. The proof method that we use here follows the same principles as those of the previous chapters. We use multivariate generating functions and analytic methods. Interestingly these methods can be applied to unlabelled rooted trees, too, where we obtain the same approximation result. And the only successful approach to the latter class of trees – also called Pólya trees – is based on generating functions in combination with Pólya's theory of counting. Thus, Pólya trees look like Galton-Watson trees although they are definitely not of that kind.

Chapter 5 considers again Galton-Watson trees but a different kind of profile that is induced by a random walk on the tree. We fix an integer valued distribution η with zero mean. Then, given a tree T, every edge e of T is endowed with an independent copy η_e of η. The label of a node is then defined as the sum of η_e over all edges e on the path to the root. There are several motivations to study such random models. For example, if η has only values ± 1 or 0 and ± 1 then the resulting trees are closely related to random triangulations and quadrangulations. Furthermore, the random variables η_e can be seen as random increments in an embedding of the tree in the space. This idea is originally due to Aldous [5] and gave rise of the ISE, the integrated super-Brownian excursion, which acts as the limiting occupation measure of the induced label distribution. The final result is that the corresponding profile can be approximated by the (random) density of the ISE. This result reaches very far and is out of scope of this book but, nevertheless, there are special cases which are of particular interest and capable for the framework of the present book. By the use of explicit generating functions of unexpected form the analysis recovers one-dimensional versions of the functional limit theorem and also leads to integral representations for several parameters of the ISE. These observations are due to Bousquet-Mélou [23].

Chapter 6 deals with recursive trees and their variants (plane oriented recursive trees, binary and m-ary search trees). The interesting feature of these kinds of trees is that they can be seen from different points of views: They can be seen as a combinatorial object (where usual counting procedures apply) as well as the result of a (stochastic) growth process. Interestingly their asymptotic structure is completely different from that of Galton-Watson trees. They are so-called $\log n$ trees which means that their expected height is of order $\log n$ (in contrast to Galton-Watson trees with expected height of order $\sqrt{n}$). We provide a unified approach to several basic statistics like the degree distribution. However, the main focus is again the profile. Here one observes that most vertices are concentrated around few levels so that a (possible) limiting object of the normalised project is not related to some functional of the Brownian motion. Nevertheless, the normalised profile $X_{n,k}/\mathbb{E}\, X_{n,k}$ can be approximated by $X(k/\log n)$, where $X(t)$ is now a random analytic function. We also deal with the height and its concentration properties.

Tries and digital search trees are two other classes of $\log n$ trees which are discussed in Chapter 7. Their construction is based on digital keys and not on the order structure of the keys as in the case of binary search trees. Again, most vertices are concentrated around few levels of order $\log n$ but the profile behaves differently. It is even more concentrated around its mean value than the profile of binary search trees or recursive trees. The normalised profile $X_{n,k}/\mathbb{E}\, X_{n,k}$ (of tries) converges to 1 and we observe a central limit theorem.

Chapter 8 is devoted to the so-called contraction method which was developed to handle stochastic recurrence relations which naturally appear in the stochastic analysis of recursive algorithms like Quicksort. Such recurrences also appear in the analysis of the profile of recursive trees and binary search trees (and their variants). The idea is that after normalisation the recurrence relation stabilises to a (stochastic) fixed point equation that can be solved uniquely by Banach's fixed point theorem in a properly chosen Banach space setting. Here we restrict ourselves to an L_2 setting with the Wasserstein metric. We mainly follow the work by Rösler, Rüschendorf, Neininger [158, 161, 162, 186, 187].

The final Chapter 9 deals with planar graphs. At first sight planar graphs and trees have nothing in common but there are strong similarities in the combinatorial and asymptotic analysis. For example the 2-connected parts of a connected (planar) graph have a tree structure which is reflected by the structure of the corresponding generating functions. In particular in the asymptotic analysis one can use the same techniques from Chapter 2 as for combinatorial tree classes in Chapter 3. Besides the asymptotic counting problem the major goal of this chapter is to study the degree distribution of random planar graphs or equivalently the expected number of vertices of given degree where we can again use asymptotic tree counting techniques. This chapter is based on recent work by Giménez, Noy and the author [63, 64].

Of course, such a book project cannot be completed without help and support from many colleagues and friends. In particular I am grateful to Mireille Bousquet-Mélou, Luc Devroye, Philippe Flajolet, Bernhard Gittenberger, Alexander Iksanov, Svante Janson, Christian Krattenthaler, Jean-François Marckert, Marc Noy, Ralph Neininger, Alois Panholzer, and Wojciech Szpankowski. I also thank Frank Emmert-Streib for helping me to design the book cover.

Finally I want to thank Veronika Kraus, Johannes Morgenbesser, and Christoph Strolz for their careful reading of the manuscript and for several hints to improve the presentation and Barbara Doležal-Rainer for her support in type setting. I also want to thank Stephen Soehnlen from Springer Verlag for his constant support in this book project and his patience.

I am especially indebted to my family to whom this book is dedicated.

Vienna, November 2008 *Michael Drmota*

Contents

1

Classes of Random Trees

In this first chapter we survey several types of random trees. We start with basic notions on trees and the description of several concepts of tree counting problems. In particular we distinguish between rooted and unrooted, plane and non-plane, and labelled and unlabelled trees. It is also possible to modify the counting procedure by putting certain weights on trees, for example, by using the degree distribution.

We consider classical combinatorial tree classes like planted plane trees or labelled rooted trees. Furthermore we discuss simply generated trees which can be also considered as conditioned Galton-Watson trees and cover several classes of the classical (rooted) trees. We introduce unlabelled trees (also called Pólya trees) that do not fall into this class but behave similarly to simply generated trees. Recursive trees (and more generally increasing trees) are labelled rooted trees where each path starting at the root has increasing labels. All these kinds of trees give rise to a natural probability distribution based on combinatorics by assuming that every tree of size n (of a certain class) is equally likely.

Trees occur also in the context of algorithms from computer science, for example, as data structures. Here the structure of the tree is determined by the input data of the algorithm. Prominent examples are binary search trees, digital search trees or tries. From a combinatorial point of view these kinds of trees are just binary trees. However, if we assume some probability distribution on the input data this induces a probability distribution on the corresponding trees. Moreover, one usually has a tree evolution process by inserting more and more data.

1.1 Basic Notions

Trees are defined as connected graphs without cycles, and their properties are basics of graph theory. For example, a connected graph is a tree, if and only if the number of edges equals the number of nodes minus 1. Furthermore, each pair of nodes is connected by a unique path.

The degree $d(v)$ of a node v in a tree is the number of nodes that are adjacent to v or the number of neighbours of v.

Nodes of degree ≤ 1 are usually called *leaves* or *external nodes* and the remaining ones *internal nodes*.

1.1.1 Rooted Versus Unrooted trees

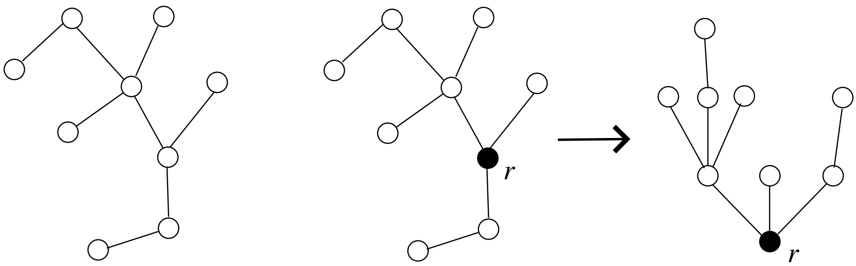

Fig. 1.1. Tree and rooted tree

If we mark a specific node r in a tree T, which we denote the *root* of T, we call the tree itself *rooted tree*. A rooted tree may be described easily in terms of *generations* or *levels*. The root is the 0-th generation. The neighbours of the root constitute the first generation, and in general the nodes at distance k from the root form the k-th generation (or level). If a node of level k has neighbours of level $k+1$ then these neighbours are also called *successors*. The number of successors of a node v is also called the *out-degree* $d^+(v)$. For all nodes v different from the root we have $d(v) = d^+(v) + 1$.

Furthermore, if v is a node in a rooted tree T then v may be considered as the root of a subtree T_v of T that consists of all iterated successors of v. This means that rooted trees can be constructed in a recursive way. Due to that property counting problems on rooted trees are usually easier than on unrooted trees.

Remark 1.1 *Rooted trees also have various applications in computer science. They naturally appear as data structures, e.g. the recursive structure of folders in any computer is just a rooted tree. Furthermore, fundamental algorithms such as Quicksort or the Lempel-Ziv data compression algorithm are closely* *related to rooted trees, namely to binary and digital search trees which are also used to store (and search for) data. Rooted trees even occur in information theory. For example, prefix free codes on an alphabet of order m are encoded as the set of leaves in m-ary trees.*

1.1.2 Plane Versus Non-Plane trees

Trees are planar graphs since they can be embedded into the plane without crossings. Nevertheless, a tree may have different embeddings (compare with Figure 1.2). This makes a difference in counting problems. When we say that we are counting *planar trees* we mean that we are counting all possible different embeddings into the plane.

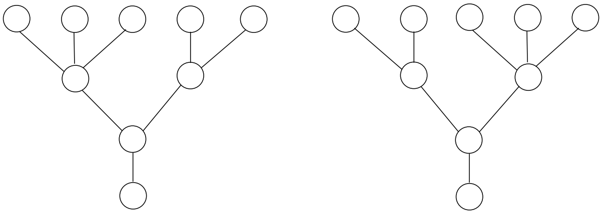

Fig. 1.2. Two different embeddings of a tree

In the context of rooted trees it is common to use the term *plane tree* or *ordered tree* when successors of the root and recursively the successors of each node are equipped with a *left-to-right-order*. Alternatively one can give the successors a rank so that one can speak of the j-th successor ($j \geq 1$). Of course, this induces a natural embedding into the half-plane (compare with Figure 1.3). Note that this notion is different from considering all embeddings into the plane, since it is not allowed to rotate the subtrees of the root cyclically around the root.

1.1.3 Labelled Versus Unlabelled Trees

We also distinguish between labelled trees, where the nodes are labelled by different numbers, and unlabelled trees, where nodes are indistinguishable. This is particularly important for the counting problem. For example, there is only one unlabelled tree with three nodes whereas there are three different labelled trees of size 3 with labels $1, 2, 3$ (see Figure 1.4).

There is much latitude in choosing labels on trees. The simplest model is to assume that the nodes of a trees of size n are labelled by the numbers $1, 2, \ldots, n$, but there are many other ways to do so. For so-called embedded trees one only assumes that the labels of adjacent vertices differ (at most) by

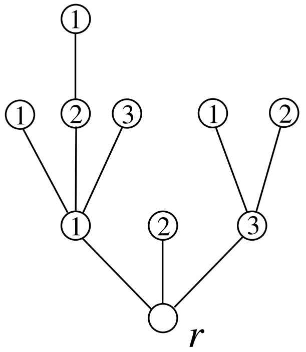

Fig. 1.3. Plane rooted tree

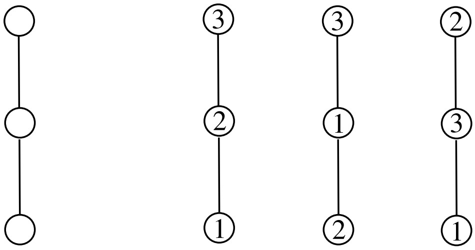

Fig. 1.4. Unlabelled versus labelled trees

1. Another possibility is to put labels consistently with the structure of the tree. For example, recursive trees have the property that the root is labelled by 1 and the labels on all paths away from the root are strictly increasing.

1.2 Combinatorial Trees

Let $\mathcal{T}$ be a class of finite trees which is defined by a structural condition (for example that the trees are binary). We then consider the subclasses $\mathcal{T}_n$ of $\mathcal{T}$ that consist of trees of size n and introduce a probability model on $\mathcal{T}_n$ by assuming that every tree T in $\mathcal{T}_n$ is equally likely. By this construction we get special kinds of random trees. Moreover, every parameter on trees (such as the number of leaves or the diameter) is then a random variable.

For simplicity we start with rooted trees since they have a recursive description.

1.2.1 Binary Trees

Binary trees are rooted trees, where each node is either a leaf (that is, it has no successor) or it has two successors. Usually these two successors are distinguishable: the left successor and the right successor, that is, we are dealing with plane trees. The leaves of a binary tree are also called *external* nodes and those nodes with two successors *internal* nodes. It is clear that a binary tree with n internal nodes has $n+1$ external nodes. Thus, the total number of nodes is always odd.

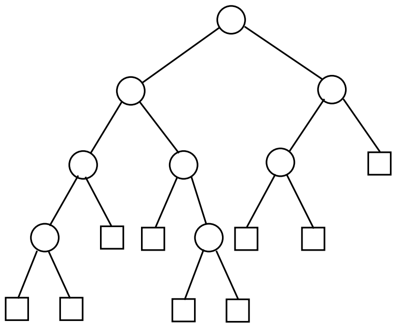

Fig. 1.5. Binary tree

A very important issue is that binary trees (and many other kinds of rooted trees) have a recursive structure. More precisely we can use the following *recursive definition* of binary trees:

> A binary tree $\mathcal{B}$ is either just an external node or an internal node (the root) with two subtrees that are again binary trees.

Formally we can write this in the form

$$\mathcal{B} = \Box + \circ \times \mathcal{B} \times \mathcal{B}, \tag{1.1}$$

where $\mathcal{B}$ denotes the system of binary trees; $\Box$ represents an external and $\circ$ an internal node.

In fact, this recursive description is the key for the analysis of many properties of binary (and similarly defined) trees. In particular, this formal equation has a direct translation into an equation for the corresponding generating (or counting) function $b(x)$ of the form $b(x) = 1 + xb(x)^2$. We discuss this translation in detail in Chapter 2.

A direct generalisation of binary trees is m-ary rooted trees, where $m \geq 2$ is a fixed integer. As in the binary case ($m = 2$) we just take into account the number n of internal nodes. The number of leaves is then given by $(m-1)n+1$ and the total number of nodes by $mn+1$.

Interestingly it is relatively easy to find explicit formulas for the numbers $b_n^{(m)}$ of m-ary trees with n internal nodes:

$$b_n^{(m)} = \frac{1}{(m-1)n+1}\binom{mn}{n}.$$

The set $\mathcal{T}_n$ of m-ary trees with n internal nodes then constitutes a set of *random trees* if we assume that every m-ary tree in $\mathcal{T}_n$ is equally likely, namely of probability $1/b_n^{(m)}$.

Note that in the binary case the number of trees is precisely the n-th *Catalan number*

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

It is also possible to consider binary and more generally m-ary trees, where the left-to-right-order of the successors is not taken into account. However, the counting problem of these classes of trees is much more involved (compare with Sections 1.2.5 and 3.1.5).

1.2.2 Planted Plane Trees

Another interesting class of trees are *planted plane trees*. Sometimes they are also called *Catalan trees*. Planted plane trees are again rooted trees, where each node has an arbitrary number of successors with a natural left-to-right-order (this again means that we are considering plane trees). The term *planted* comes from the interpretation that the root is connected (or planted) to an additional *phantom* node that is not taken into account (see Figure 1.6). Usually we will not even depict this additional node when we deal with planted trees. However, it is quite useful to define the degree of the root r by $d(r) = d^+(r) + 1$ which means that the additional (planted) node is considered a neighbour node. This has the advantage that in this case all nodes have the property $d(v) = d^+(v) + 1$.

The numbers p_n of planted plane trees with $n \geq 1$ nodes are given by

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

This is precisely the $(n-1)$-st Catalan number C_{n-1} which explains the term *Catalan tree*. By the way, the relation $p_{n+1} = b_n$ has a natural interpretation (see Section 3.1.2).

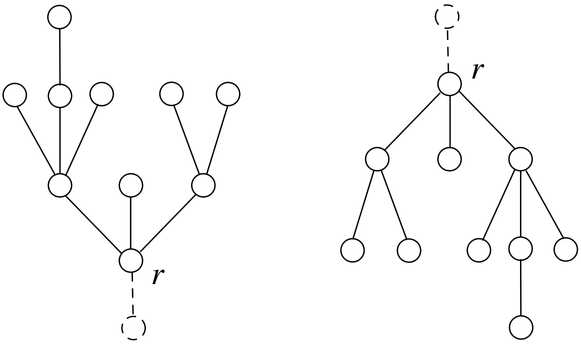

Fig. 1.6. Planted plane tree

1.2.3 Labelled Trees

We recall that a tree T of size n is labelled if the n nodes are labelled by $1, 2, \ldots, n$.[1] The counting problem of labelled trees is different from that of unlabelled trees. There is, however, an easy connection between rooted and unrooted labelled trees. There are exactly n different ways to make an unrooted tree to a rooted one by choosing one of the labelled nodes. Thus, the number of rooted labelled trees of size n equals the number of unrooted labelled trees exactly n times. Consequently it is sufficient to consider rooted labelled trees which has the advantage that one can use the recursive structure.

Note that if we do not care about the embedding in the plane or about the left to right order of the successors, an unrooted labelled tree can be interpreted as a spanning tree of the complete graph K_n with nodes $1, 2, \ldots, n$ (see Figure 1.7).

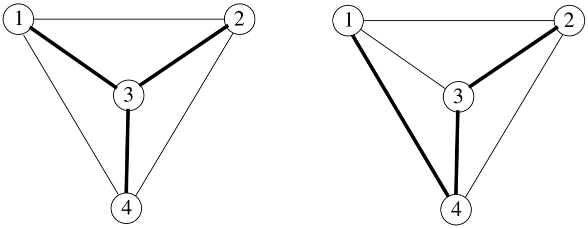

Fig. 1.7. 2 of 16 possible spanning trees of K_4

[1] Other kinds of labelled trees like recursive trees or well-labelled trees will be discussed in the sequel.

It is a well known fact that the number of unrooted labelled trees of size n equals n^{n-2} (usually called Cayley's formula). Hence, there are n^{n-1} different rooted labelled trees of size n. Sometimes these trees are called *Cayley trees* (but this term is also used for infinite regular trees).

1.2.4 Labelled Plane Trees

It is also of interest to count the number of different planar embeddings of labelled trees. There is even an explicit formula, namely for $n \ge 2$ there are

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

different planar embeddings of labelled trees of size n (and $n(2n-3)!/(n-1)!$ different planar embeddings of rooted labelled trees of size n). For example, for $n = 4$ there are $4^2 = 16$ different labelled trees but $5!/3! = 20$ different planar embeddings.

1.2.5 Unlabelled Trees

Let $\tilde{\mathcal{T}}$ denote the set of unlabelled unrooted trees and $\mathcal{T}$ be the set of unlabelled rooted trees. Here we do not care about the possible embeddings into the plane. We just think of trees in the graph-theoretical sense.

These kinds of trees are relatively difficult to count. Let us denote by $\tilde{t}_n$ and t_n the corresponding numbers of those trees of size n, for example we have

$$\tilde{t}_1 = 1,\ \tilde{t}_2 = 1,\ \tilde{t}_3 = 1,\ \tilde{t}_4 = 2 \quad \text{and} \quad t_1 = 1,\ t_2 = 1,\ t_3 = 2,\ t_4 = 4.$$

However, if there is no direct recursive relation one has to take into account all symmetries. Nevertheless, this problem can be solved by using generating functions and Pólya's theory of counting [176] (see Section 3.1.5). For that reason these trees are also called *Pólya trees*.

In order to give an impression of the kind of problems one has to face we just state that the generating functions

$$\tilde{t}(x) = \sum_{n\ge1} \tilde{t}_n x^n \quad \text{and} \quad t(x) = \sum_{n\ge1} t_n x^n$$

satisfy the relations

$$t(x) = x \exp\left(t(x) + \frac{1}{2}t(x^2) + \frac{1}{3}t(x^3) + \cdots \right) \tag{1.2}$$

and

$$\tilde{t}(x) = t(x) - \frac{1}{2}t(x)^2 + \frac{1}{2}t(x^2). \tag{1.3}$$

It seems that there is no proper explicit formula for t_n and $\tilde{t}_n$. However, there are asymptotic expansions for them and by using extensions of the mentioned counting procedure it is also possible to study several shape characteristics of these kinds of trees.

1.2.6 Unlabelled Plane Trees

We already mentioned that a tree usually has several different embeddings into the plane. Planted plane trees are, in particular, designed to take into account all possible planar embeddings of planted rooted trees.

It is, however, another non-trivial step to count all embeddings of unlabelled rooted trees and all embeddings of unlabelled trees. Again we have to take into account symmetries. Fortunately Pólya's theory can be applied here, too. As in the case of unlabelled trees we do not get explicit formulas but asymptotic expansions (see Section 3.1.6).

1.2.7 Simply Generated Trees – Galton-Watson Trees

Simply generated trees are weighted versions of rooted trees and have been introduced by Meir and Moon [151]. The idea is to put a weight to a rooted tree according to its degree distribution.

Let ϕ_j, $j \geq 0$, be a sequence of non-negative real numbers, called the *weight sequence*. Usually one assumes that $\phi_0 > 0$ and $\phi_j > 0$ for some $j \geq 2$. We then define the weight $\omega(T)$ of a finite rooted ordered tree T by

$$\omega(T) = \prod_{v \in V(T)} \phi_{d^+(v)} = \prod_{j \geq 0} \phi_j^{D_j(T)},$$

where $d^+(v)$ denotes the out-degree of the vertex v (or the number of successors) and $D_j(T)$ the number of nodes in T with j successors. The numbers

$$y_n = \sum_{|T|=n} \omega(T)$$

are then the weighted numbers of trees of size n. It is natural to define a probability distribution on the set $\mathcal{T}_n$ by

$$\pi_n(T) = \frac{\omega(T)}{y_n} \qquad (T \in \mathcal{T}_n). \tag{1.4}$$

It is convenient to introduce the generating series

$$\Phi(x) = \phi_0 + \phi_1 x + \phi_2 x^2 + \cdots = \sum_{j \geq 0} \phi_j x^j.$$

In Section 3.1.4 we will show that the generating function $y(x) = \sum_{n \geq 1} y_n x^n$ satisfies the equation

$$y(x) = x\,\Phi(y(x)).$$

This equation is the key for the asymptotic analysis of these kinds of trees.

If we replace ϕ_j by $\tilde{\phi}_j = ab^j \phi_j$, which is the same as replacing $\Phi(x)$ by $\tilde{\Phi}(x) = a\Phi(bx)$ for two numbers $a, b > 0$, then $\omega(T)$ is replaced by

$$\tilde{\omega}(T) = \prod_{j \ge 0} \left(ab^j \phi_j\right)^{D_j(T)} = a^{|T|} b^{|T|-1} \omega(T).$$

Note that $\sum_j jD_j(T) = |T| - 1$. Hence, $\tilde{y}_n = a^n b^{n-1} y_n$ and the probability distribution π_n on $\mathcal{T}_n$ is the same for $\tilde{\Phi}(x)$ and $\Phi(x)$ (for every n). Usually only these distributions are important, and we may then freely make this type of modification of ϕ_j.

Simply generated trees generalise several of the above examples of combinatorial trees.

Example 1.2 *If $\phi_j = 1$ for all $j \ge 0$, that is, $\Phi(x) = 1/(1-x)$, then all planted plane trees have weight $\omega(T) = 1$ and y_n is the number of planted plane trees. Thus, π_n is the uniform distribution on planted plane trees of size n.*

Example 1.3 *Binary trees (counted according to their internal nodes) are also covered by this approach. If we set $\phi_0 = 1$, $\phi_1 = 2$, $\phi_2 = 1$, and $\phi_j = 0$ for $j \ge 3$, that is, $\Phi(x) = (1+x)^2$, then nodes with one successor get weight 2. This takes into account that binary trees (where external nodes are disregarded) have two kinds of nodes with one successor, namely those with a left branch but no right branch and those with a right branch but no left branch. Thus, π_n is the uniform distribution on all binary trees with n internal nodes.*

Similarly, m-ary trees are covered with the help of the weights $\phi_j = \binom{m}{j}$ or with $\Phi(x) = (1+x)^m$.

Example 1.4 *If $\phi_0 = \phi_1 = \phi_2 = 1$ and $\phi_j = 0$ for $j \ge 3$ or $\Phi(x) = 1+x+x^2$, then we get so-called Motzkin trees. Here only rooted trees, where all nodes have less than 3 successors, get (a non-zero) weight $\omega(T) = 1$: y_n is the number of Motzkin trees with n nodes and π_n is the uniform distribution on Motzkin trees of size n.*

Example 1.5 *If we set $\phi_j = 1/j!$ then*

$$n! \cdot y_n = n^{n-1}$$

denotes precisely the number of labelled rooted non-plane trees. The weight $\phi_j = 1/j!$ disregards all possible orderings of the successors of a vertex of out-degree j and the factor $n!$ corresponds to all possible labellings of n nodes. Hence, π_n yields the uniform distribution on labelled rooted trees.

Interestingly there is an intimate relation to *Galton-Watson branching processes*. Let ξ be a non-negative integer-valued random variable, the so-called offspring distribution. The Galton-Watson branching process starts with a single individual (generation 0); each individual has a number of children distributed as independent copies of ξ. If Z_k denotes the size of the generation k, then a formal description of the process $(Z_k)_{k \ge 0}$ is $Z_0 = 1$, and for $k \ge 1$

$$Z_k = \sum_{j=1}^{Z_{k-1}} \xi_j^{(k)},$$

where the $(\xi_j^{(k)})_{k,j}$ are i.i.d.[2] random variables distributed as ξ.

It is clear that Galton-Watson branching processes can be represented by ordered (finite or infinite) rooted trees T such that the sequence Z_k is just the number of nodes at level k and $\sum_{k\ge 0} Z_k$ (which is called the *total progeny*) is the number of nodes $|T|$ of T. We denote by $\nu(T)$ the probability that a specific tree T occurs. If $\mathbb{P}\{\xi = 0\} = 0$ then the total progeny is infinite with probability 1. Thus we always assume that $\mathbb{P}\{\xi = 0\} > 0$.

The generating function $y(x) = \sum_{n\ge 1} y_n x^n$ of the numbers

$$y_n = \mathbb{P}\{|T| = n\} = \sum_{|T|=n} \nu(T)$$

satisfies the functional equation

$$y(x) = x\,\Phi(y(x)),$$

where

$$\Phi(t) = \mathbb{E}\, t^\xi = \sum_{j\ge 0} \phi_j t^j$$

with $\phi_j = \mathbb{P}\{\xi = j\}$. Observe that

$$\nu(T) = \prod_{j\ge 0} \phi_j^{D_j(T)} = \omega(T).$$

The *weight* of T is now the *probability* of T.

If we condition the Galton-Watson tree T on $|T| = n$, we thus get the probability distribution (1.4) on $\mathcal{T}_n$. Hence, the conditioned Galton-Watson trees are simply generated trees with $\phi_j = \mathbb{P}\{\xi = j\}$ as above. We have here $\Phi(1) = \sum_j \phi_j = 1$, but this is no real restriction. In fact, if $(\phi_j)_{j\ge 0}$ is any sequence of non-negative weights satisfying the very weak condition $\Phi(x) = \sum_{j\ge 0} \phi_j x^j < \infty$ for some $x > 0$, then we can replace (as above) ϕ_j by $ab^j\phi_j$ with $b = x$ and $a = 1/\Phi(x)$ and thus the simply generated tree is the same as the conditioned Galton-Watson tree with offspring distribution $\mathbb{P}\{\xi = j\} = \phi_j x^j/\Phi(x)$. Consequently, for all practical purposes, simply generated trees are the same as conditioned Galton-Watson trees.

The argument above also shows that the distribution of a conditioned Galton-Watson tree is not changed if we replace the offspring distribution ξ by $\tilde{\xi}$ with $\mathbb{P}\{\tilde{\xi} = j\} = \mathbb{P}\{\xi = j\} = \tau^j/\Phi(\tau)$ and thus $\tilde{\Phi}(x) = \Phi(\tau x)/\Phi(\tau)$ for any $\tau > 0$ with $\Phi(\tau) < \infty$. (Such modifications are called conjugate or tilted distributions.)

[2] The letters "i.i.d." abbreviate "independent and identically distributed".

Note that

$$\mu = \Phi'(1) = \mathbb{E}\,\xi$$

is the expected value of the offspring distribution. If $\mu < 1$, the Galton-Watson branching process is called sub-critical, if $\mu = 1$, then it is critical, and if $\mu > 1$, then it is supercritical. From a combinatorial point of view we do not have to distinguish between these three cases. Namely, if we replace the offspring distribution by a conjugate distribution as above, the new expected value is

$$\tilde{\Phi}'(1) = \frac{\tau\Phi'(\tau)}{\Phi(\tau)}.$$

We can thus always assume that the Galton-Watson process is critical, provided only that there exists $\tau > 0$ with

$$\tau\Phi'(\tau) = \Phi(\tau) < \infty,$$

a weak condition that is satisfied for all interesting classes of Galton-Watson trees.

It is usually convenient to choose a critical version, which explains why the equation $\tau\Phi'(\tau) = \Phi(\tau)$ appears in most asymptotic results. A heuristic reason is that the probability of the event $|T| = n$ that we condition on typically decays exponentially in the subcritical and supercritical cases but only as $n^{-1/2}$ in the critical case, and it seems advantageous to condition on an event of not too small probability.

Example 1.6 *For planted plane trees (as in Example 1.2) we start with* $\Phi(x) = 1/(1-x)$. *The equation* $\tau\Phi'(\tau) = \Phi(\tau)$ *is* $\tau(1-\tau)^{-2} = (1-\tau)^{-1}$, *which is solved by* $\tau = \frac{1}{2}$. *Random planted plane trees are thus conditioned Galton-Watson trees with the critical offspring distribution given by* $\Phi(x) = (1-x/2)^{-1}/2 = 1/(2-x)$, *or* $\mathbb{P}\{\xi = j\} = 2^{-j-1}$ *(for* $j \ge 0$*), a geometric distribution.*

Example 1.7 *Similarly random binary trees are obtained with a binomial offspring distribution* $\mathrm{Bi}(2, 1/2)$ *with* $\Phi(x) = (1+x)^2/4$, *and more generally random* m*-ary trees are obtained with offspring law* $\mathrm{Bi}(m, 1/m)$ *with* $\Phi(x) = ((m-1+x)/m)^m$.

Example 1.8 *For Motzkin trees the critical offspring distribution* ξ *is uniform on* $\{0, 1, 2\}$ *with* $\Phi(x) = (1+x+x^2)/3$.

Example 1.9 *For uniform rooted labelled trees the critical* ξ *has a Poisson distribution* $\mathrm{Po}(1)$ *with* $\Phi(x) = e^{x-1}$.

Finally we remark that for a critical offspring distribution ξ, its variance is given by

$$\sigma^2 = \mathbb{V}\mathrm{ar}\,\xi = \mathbb{E}\,\xi^2 - 1 = \mathbb{E}(\xi(\xi-1)) = \Phi''(1).$$

Starting with an arbitrary sequence $(\phi_j)_{j\geq 0}$ and modifying it as above the get a critical probability distribution, we obtain the variance

$$\sigma^2 = \tilde{\Phi}''(1) = \frac{\tau^2 \Phi''(\tau)}{\Phi(\tau)},$$

where $\tau > 0$ is such that $\tau\Phi'(\tau) = \Phi(\tau) < \infty$ (assuming this is possible). We will see that this quantity appears in several asymptotic results.

1.3 Recursive Trees

Recursive trees are rooted labelled trees, where the root is labelled by 1 and the labels of all successors of any node v are larger than the label of v (see Figure 1.8).

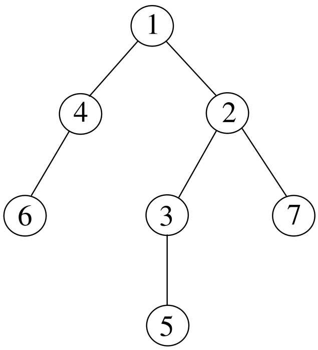

Fig. 1.8. Recursive tree

1.3.1 Non-Plane Recursive Trees

Usually one does not take care of the possible embeddings of a recursive tree into the plane. In this sense recursive trees can be seen as the result of the following evolution process. Suppose that the process starts with a node carrying the label 1. This node will be the root of the tree. Then attach a node with label 2 to the root. The next step is to attach a node with label 3. However, there are two possibilities: either to attach it to the root or to the node with label 2. Similarly one proceeds further. After having attached the nodes with labels $1, 2, \ldots, k$, attach the node with label $k + 1$ to one of the existing nodes.

Obviously, every recursive tree of size n is obtained in a unique way. Moreover, the labels represent something like the history of the evolution process.

Since there are exactly k ways to attach the node with label $k+1$, there are exactly $(n-1)!$ possible trees of size n.

The natural probability distribution on recursive trees of size n is to assume that each of these $(n-1)!$ trees is equally likely. This probability distribution is also obtained from the evolution process by attaching successively each new node to one of the already existing nodes with equal probability.

Remark 1.10 *Historically, recursive trees appear in various contexts. They are used to model the spread of epidemics (see [155]) or to investigate and construct family trees of preserved copies of ancient manuscripts (see [157]). Other applications are the study of the schemes of chain letters or pyramid games (see [88]).*

1.3.2 Plane Oriented Recursive Trees

Note that the left-to-right-order of the successors of the nodes in a recursive tree was not relevant in the above counting procedure. It is, however, relatively easy to consider all possible embeddings as plane rooted trees. These kind of trees are usually called plane oriented recursive trees (PORTs).

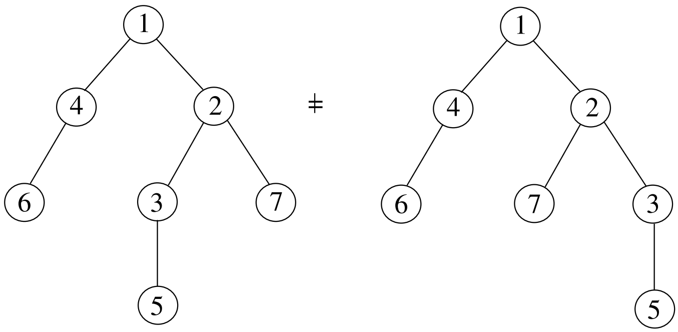

Fig. 1.9. Two different plane oriented trees

They can again be seen as the result of an evolution process, where the left-to-right-order of the successors is taken into account. More precisely, if a node v has out-degree d, then there are $d+1$ possible ways to attach a new node to v. Hence, the number of different plane oriented recursive trees with n nodes equals

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

As above, the natural probability distribution on plane oriented recursive trees of size n is to assume that each of these $(2n-3)!!$ trees is equally likely.

This probability distribution is also obtained from the evolution process by attaching each node with probability proportional to the out-degree plus 1 to the already existing nodes.

1.3.3 Increasing Trees

The probabilistic model of simply generated trees was to define a weight that reflects the degree distribution of rooted trees. The same idea can be applied to recursive and to plane oriented recursive trees. The resulting classes of trees are called *increasing trees*. They have been first introduced by Bergeron, Flajolet, and Salvy [12].

As above we define the weight $\omega(T)$ of a recursive or a plane oriented recursive tree T by

$$\omega(T) = \prod_{v \in V(T)} \phi_{d^+(v)} = \prod_{j \geq 0} \phi_j^{D_j(T)},$$

where $d^+(v)$ denotes the out-degree of the vertex v (or the number of successors) and $D_j(T)$ the number of nodes in T with j successors. Then we set

$$y_n = \sum_{T \in \mathcal{J}_n} \omega(T),$$

where $\mathcal{J}_n$ denotes the set of recursive or plane oriented recursive trees of size n. The natural probability distribution on the set $\mathcal{J}_n$ of increasing trees is then given by

$$\pi_n(T) = \frac{\omega(T)}{y_n} \qquad (T \in \mathcal{J}_n).$$

As in the case of simply generated trees it is also possible to introduce generating series. We set

$$\Phi(x) = \phi_0 + \phi_1 x + \phi_2 \frac{x^2}{2!} + \phi_2 \frac{x^3}{3!} + \cdots$$

in the case of recursive trees and

$$\Phi(x) = \phi_0 + \phi_1 x + \phi_2 x^2 + \phi_3 x^3 + \cdots$$

in the case of plane oriented recursive trees. The generating function

$$y(z) = \sum_{n \geq 0} y_n \frac{z^n}{n!}$$

satisfies the differential equation

$$y'(z) = \Phi(y(z)), \quad y(0) = 0.$$

In the interest of clarity we state how the general concept specialises.

1. Recursive trees (that is, every non-planar recursive tree gets weight 1) are given by $\Phi(x) = e^x$. Here $y_n = (n-1)!$ and $y(z) = \log(1/(1-z))$.
2. Plane oriented recursive trees are given by $\Phi(x) = 1/(1-x)$. This means that every planar recursive tree gets weight 1. Here $y_n = (2n-3)!! = 1 \cdot 3 \cdot 5 \cdots (2n-3)$ and $y(z) = 1 - \sqrt{1-2z}$.
3. Binary recursive trees are defined by $\Phi(x) = (1+x)^2$. We have $y_n = n!$ and $y(z) = 1/(1-z)$. The probability model that is induced by this (planar) binary increasing trees is exactly the standard permutation model of *binary search trees* that is discussed in Section 1.4.1.

Note that the probability distribution on $\mathcal{J}_n$ is not automatically given by an evolution process as it is definitely the case for recursive trees and plane oriented recursive trees. It is interesting that there are precisely three families of increasing trees, where the probability distribution π_n is also induced by a (natural) tree evolution process.

1. $\Phi(x) = \phi_0 e^{\frac{\phi_1}{\phi_0}x}$ with $\phi_0 > 0$, $\phi_1 > 0$.
2. $\Phi(x) = \phi_0 \left(1 - \dfrac{\phi_1}{r\phi_0}x\right)^{-r}$ for some $r > 0$ and $\phi_0 > 0$, $\phi_1 > 0$.
3. $\Phi(x) = \phi_0 \left(1 + (\phi_1/(d\phi_0))x\right)^d$ for some $d \in \{2, 3, \ldots\}$ and $\phi_0 > 0$, $\phi_1 > 0$.

The corresponding tree evolution process runs as follows:[3] The starting point is (again) a node (the root) with label 1. Now assume that a tree T of size n is present. We attach to every node v of T a *local weight* $\rho(v) = (k+1)\phi_{k+1}\phi_0/\phi_k$ when v has k successors and set $\rho(T) = \sum_{v \in V(T)} \rho(v)$. Observe that in a planar tree there are $k+1$ different ways to attach a new (labelled) node to an (already existing) node with k successors. Now choose a node v in T according to the probability distribution $\rho(v)/\rho(T)$ and then independently and uniformly one of the $k+1$ possibilities to attach a new node there (when v has k successors). This construction ensures that in these three particular cases a tree T of size n, which occurs with probability proportional to $\omega(T)$, generates a tree T' of size $n+1$ with probability that is proportional to $\omega(T)\phi_{k+1}\phi_0/\phi_k$, which equals $\omega(T')$. Thus, this procedure induces the same probability distribution on $\mathcal{J}_n$ as the one mentioned above, where a tree $T \in \mathcal{J}_n$ has probability $\omega(T)/y_n$.

Note that if we are only interested in the distributions π_n, then we can work (without loss of generality) with some special values for ϕ_0 and ϕ_1. It is sufficient to consider the generating functions

1. $\Phi(x) = e^x$,
2. $\Phi(x) = (1-x)^{-r}$ for some $r > 0$,
3. $\Phi(x) = (1+x)^d$ for some $d \in \{2, 3, \ldots\}$.

The first class is just the class of recursive trees. The second class can be interpreted as generalised plane oriented recursive trees, since the probability

[3] In the interest of brevity we only discuss the plane version.