WEIGHTED UNARY-BINARY TREES, HEX-TREES, MARKED ORDERED TREES, AND RELATED STRUCTURES

HELMUT PRODINGER

ABSTRACT. Hex-trees are identified as a particular instance of weighted unary-binary trees. The Horton-Strahler numbers of these objects are revisited, and, thanks to a substitution that is not immediately intuitive, explicit results are possible. They are augmented by asymptotic evaluations as well. Furthermore, marked ordered trees (in bijection to skew Dyck paths) are investigated, followed 3-Motzkin paths and multi-edge trees. The underlying theme is sequence A002212 in the Encyclopedia of integer sequences.

1. INTRODUCTION

This paper gives some (mostly new) results about the sequence

$$1, 1, 3, 10, 36, 137, 543, 2219, 9285, 39587, 171369, 751236, 3328218, 14878455, \dots,$$

which is A002212 in [12].

Here is the plan about the structures enumerated by this sequence:

Hex-trees [9]; they are identified as weighted unary-binary trees, with weight one. Apart from left and right branches, as in binary trees, there are also unary branches, and they can come in different colours, here in just one colour. Unary-binary trees played a role in the present authors scientific development, as documented in [6], a paper written with the late and great Philippe Flajolet, about the register function (Horton-Strahler numbers) of unary-binary trees. Here, we can offer an improvement, using a "better" substitution than in [6]. The results can now be made fully explicit. As a by-product, this provides a definition and analysis of the Horton-Strahler numbers of Hex-trees. An introductory section (about binary trees) provides all the basics.

Then we move to skew Dyck paths [2]. They are like Dyck paths, but allow for an extra step $(-1, -1)$, provided that the path does not intersect itself. An equivalent model, defined and described using a bijection, is from [2]: Marked ordered trees. They are like ordered trees, with an additional feature, namely each rightmost edge (except for one that leads to a leaf) can be coloured with two colours. Since we find this class of trees to be interesting, we analyze two parameters of them: number of leaves and height. While the number of leaves for ordered trees is about $n/2$, it is only $n/10$ in the new model. For the height, the leading term $\sqrt{\pi n}$ drops to $\frac{2}{\sqrt{5}}\sqrt{\pi n}$. Of course, many more parameters of this new class of trees could be investigated, which we encourage to do.

The last two classes of structures are multi-edge trees. Our interest in them was already triggered in an earlier publication [8]. They may be seen as ordered trees, but

with weighted edges. The weights are integers ≥ 1, and a weight a may be interpreted as a parallel edges. The other class are 3-Motzkin paths. They are like Motzkin paths (Dyck paths plus horizontal steps); however, the horizontal steps come in three different colours. A bijection is described. Since 3-Motzkin paths and multi-edge trees are very much alike (using a variation of the classical rotation correspondence), all the structures that are discussed in this paper can be linked via bijections.

2. Binary trees and Horton-Strahler numbers

Binary trees may be expressed by the following symbolic equation, which says that they include the empty tree and trees recursively built from a root followed by two subtrees, which are binary trees:

$$\mathscr{B} \quad = \quad \square \quad + \quad \underset{\mathscr{B} \quad \mathscr{B}}{\overset{\bigcirc}{\wedge}}$$

Binary trees are counted by Catalan numbers and there is an important parameter reg, which in Computer Science is called the register function. It associates to each binary tree (which is used to code an arithmetic expression, with data in the leaves and operators in the internal nodes) the minimal number of extra registers that is needed to evaluate the tree. The optimal strategy is to evaluate difficult subtrees first, and use one register to keep its value, which does not hurt, if the other subtree requires less registers. If both subtrees are equally difficult, then one more register is used, compared to the requirements of the subtrees. This natural parameter is among Combinatorialists known as the Horton-Strahler numbers, and we will adopt this name throughout this paper.

There is a recursive description of this function: $\mathsf{reg}(\square) = 0$, and if tree t has subtrees t_1 and t_2, then

$$\mathsf{reg}(t) = \begin{cases} \max\{\mathsf{reg}(t_1), \mathsf{reg}(t_2)\} & \text{if } \mathsf{reg}(t_1) \neq \mathsf{reg}(t_2), \\ 1 + \mathsf{reg}(t_1) & \text{otherwise.} \end{cases}$$

The recursive description attaches numbers to the nodes, starting with 0's at the leaves and then going up; the number appearing at the root is the Horton-Strahler number of the tree.
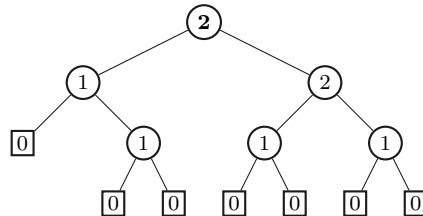
Let $\mathscr{R}_p$ denote the family of trees with Horton-Strahler number $= p$, then one gets immediately from the recursive definition:

$$\mathscr{R}_p = \underset{\mathscr{R}_{p-1} \; \mathscr{R}_{p-1}}{\overset{\bigcirc}{\wedge}} \quad + \quad \underset{\mathscr{R}_p \; \sum_{j<p} \mathscr{R}_j}{\overset{\bigcirc}{\wedge}} \quad + \quad \underset{\sum_{j<p} \mathscr{R}_j \; \mathscr{R}_p}{\overset{\bigcirc}{\wedge}}$$

In terms of generating functions, these equations read as

$$R_p(z) = zR_{p-1}^2(z) + 2zR_p(z)\sum_{j<p} R_j(z);$$

the variable z is used to mark the size (i. e., the number of internal nodes) of the binary tree.

A historic account of these concepts, from the angle of Philippe Flajolet, who was one of the pioneers is [11], compare also [10].

Amazingly, the recursion for the generating functions $R_p(z)$ can be solved explicitly! The substitution

$$z = \frac{u}{(1+u)^2}$$

that de Bruijn, Knuth, and Rice [1] also used, produces the nice expression

$$R_p(z) = \frac{1-u^2}{u}\frac{u^{2^p}}{1-u^{2^{p+1}}}.$$

Of course, once this is *known*, it can be proved by induction, using the recursive formula. For the readers benefit, this will be sketched now.

We start with the auxiliary formula

$$\sum_{0\le j<p} \frac{u^{2^j}}{1-u^{2^{j+1}}} = \frac{u}{1-u} - \frac{u^{2^p}}{1-u^{2^p}},$$

which is easy to prove by induction: For $p = 0$, the formula $0 = \frac{u}{1-u} - \frac{u}{1-u}$ is correct, and then

$$\sum_{0\le j<p+1} \frac{u^{2^j}}{1-u^{2^{j+1}}} = \frac{u}{1-u} - \frac{u^{2^p}}{1-u^{2^p}} + \frac{u^{2^p}}{1-u^{2^{p+1}}}$$

$$= \frac{u}{1-u} - \frac{u^{2^p}(1+u^{2^p})}{1-u^{2^{p+1}}} + \frac{u^{2^p}}{1-u^{2^{p+1}}} = \frac{u}{1-u} - \frac{u^{2^{p+1}}}{1-u^{2^{p+1}}}.$$

Now the formula for $R_p(z)$ can also be proved by induction. First, $R_0(z) = \frac{1-u^2}{u}\frac{u}{1-u^2} = 1$, as it should, and

$$zR_{p-1}^2(z) + 2zR_p(z)\sum_{j<p} R_j(z)$$

$$= \frac{u}{(1+u)^2}\frac{(1-u^2)^2}{u^2}\frac{u^{2^p}}{(1-u^{2^p})^2} + \frac{2u}{(1+u)^2}R_p(z)\sum_{j<p}\frac{1-u^2}{u}\frac{u^{2^j}}{1-u^{2^{j+1}}}$$

$$= \frac{u^{2^p-1}(1-u)^2}{(1-u^{2^p})^2} + \frac{2(1-u)}{(1+u)}R_p(z)\sum_{j<p}\frac{u^{2^j}}{1-u^{2^{j+1}}}.$$

Solving

$$R_p(z) = \frac{u^{2^p-1}(1-u)^2}{(1-u^{2^p})^2} + \frac{2(1-u)}{(1+u)}R_p(z)\left[\frac{u}{1-u} - \frac{u^{2^p}}{1-u^{2^p}}\right]$$

leads to

$$R_p(z)\frac{1-u}{1+u}\left[1 + 2\frac{u^{2^p}}{1-u^{2^p}}\right] = \frac{u^{2^p-1}(1-u)^2}{(1-u^{2^p})^2},$$

or, simplified

$$R_p(z) = \frac{u^{2^p-1}(1-u^2)}{(1-u^{2^p})(1+u^{2^p})} = \frac{1-u^2}{u} \frac{u^{2^p}}{(1-u^{2^{p+1}})},$$

which is the formula that we needed to prove.   □

3. Unary-binary trees and Hex-trees

The family of unary-binary trees $\mathscr{M}$ might be defined by the symbolic equation

The equation for the generating function is

$$M = 1 + z(M-1) + zM^2$$

with the solution

$$M = M(z) = \frac{1 - z - \sqrt{1 - 6z + 5z^2}}{2z} = 1 + z + 3z^2 + 10z^3 + 36z^4 + \cdots ;$$

the coefficients form again sequence A002212 in [12] and enumerate Schröder paths, among many other things. We will come to equivalent structures a bit later.

In the instance of unary-binary trees, we can also work with a substitution: Set $z = \frac{u}{1+3u+u^2}$, then $M(z) = 1 + u$. Unary-binary trees and the register function were investigated in [6], but the present favourable substitution was not used. Therefore, in this previous paper, asymptotic results were available but no explicit formulae.

This works also with a weighted version, where we allow unary edges with a different colours. Then

and with the substitution $z = \frac{u}{1+(a+2)u+u^2}$, the generating function is beautifully expressed as $N(z) = 1 + u$. For $a = 0$, this covers also binary trees.

We will consider the Horton-Strahler numbers of unary-binary trees in the sequel. The definition is naturally extended by

$$\mathsf{reg}\left(\underset{t}{\bigcirc}\right) = \mathsf{reg}(t).$$

Now we can move again to $R_p(z)$, the generating funciton of (generalized) unary-binary trees with Horton-Strahler number $= p$. The recursion (for $p \geq 1$) is

In terms of generating functions, these equations read as

$$R_p(z) = zR_{p-1}^2(z) + 2zR_p(z)\sum_{j<p} R_j(z) + azR_p(z), \quad p \geq 1; \quad R_0(z) = 1.$$

Amazingly, with the substitution $z = \frac{u}{1+(a+2)u+u^2}$, formally we get the *same* solution as in the binary case:

$$R_p(z) = \frac{1-u^2}{u} \frac{u^{2^p}}{1-u^{2^{p+1}}}.$$

The proof by induction is as before. One sees another advantage of the substitution: On a formal level, many manipulations do not need to be repeated. Only when one switches back to the z-world, things become different.

Now we move to Hex-trees.

Hex trees either have two non-empty successors, or one of 3 types of unary successors (called left, middle, right). The author has seen this family first in [9], but one can find older literature following the references and the usual search engines.

The generating function satisfies

$$H(z) = 1 + z(H(z) - 1)^2 + z + 3z(H(z) - 1) = \frac{1 - z - \sqrt{(1-z)(1-5z)}}{2z}$$

$$= 1 + z + 3z^2 + 10z^3 + 36z^4 + 137z^5 + 543z^6 + 2219z^7 + 9285z^8 + 39587z^9 + \cdots .$$

The same generating function also appears in [8], and it is again sequence A002212 in [12]. One can rewrite the symbolic equation as

and sees in this way that the Hex-trees are unary-binary trees (with parameter $a = 1$).

Continuing with enumerations. First, we will enumerate the number of (generalized) unary-binary trees with n (internal) nodes. For that we need the notion of generalized trinomial coefficients, viz.

$$\binom{n; 1, a, 1}{k} := [z^k](1 + az + z^2)^n.$$

Of course, for $a = 2$, this simplifies to a binomial coefficient $\binom{2n}{k}$. We will use contour integration to pull out coefficients, and the contour of integer, in whatever variable, is a small circle (or equivalent) around the origin. The desired number is

$$[z^n](1 + u) = \frac{1}{2\pi i} \oint \frac{dz}{z^{n+1}}(1 + u)$$

$$= \frac{1}{2\pi i} \oint \frac{du(1 - u^2)(1 + (a+2)u + u^2)^{n+1}}{(1 + (a+2)u + u^2)^2 u^{n+1}}(1 + u)$$

$$= [u^{n+1}](1 - u)(1 + u)^2(1 + (a+2)u + u^2)^{n-1}$$

$$= \binom{n-1; 1, a+2, 1}{n+1} + \binom{n-1; 1, a+2, 1}{n}$$

$$- \binom{n-1; 1, a+2, 1}{n-1} - \binom{n-1; 1, a+2, 1}{n-2}.$$

Then we introducte $S_p(z) = R_p(z) + R_{p+1}(z) + R_{p+2}(z) + \cdots$, the generating function of trees with Horton-Strahler number $\geq p$. Using the summation formula proved earlier, we get

$$S_p(z) = \frac{1-u^2}{u}\frac{u^{2^p}}{1-u^{2^p}} = \frac{1-u^2}{u}\sum_{k\geq 1} u^{k2^p}.$$

Further,

$$[z^n]S_p(z) = \sum_{k\geq 1}\frac{1}{2\pi i}\oint \frac{dz}{z^{n+1}}\frac{1-u^2}{u}u^{k2^p}.$$

Asymptotics. We start by deriving asymptotics for the number of (generalized) unary-binary trees with n (internal) nodes. This is a standard application of singularity analysis of generating functions, as described in [5] and [7].

We start from the generating function

$$N(z) = \frac{1-az-\sqrt{1-2(a+2)z+a(a+4)z^2}}{2z}$$

and determine the singularity closest to the origin, which is the value making the square root disappear: $z = \frac{1}{a+4}$. After that, the local expansion of $N(z)$ around this singularity is determined:

$$N(z) \sim 2 - \sqrt{a+4}\sqrt{1-(a+4)z}.$$

The translation lemmas given in [5] and [7] provide the asymptotics:

$$[z^n]N(z) \sim [z^n]\left(2 - \sqrt{a+4}\sqrt{1-(a+4)z}\right)$$
$$= -\sqrt{a+4}(a+4)^n\frac{n^{-3/2}}{\Gamma(-\frac{1}{2})} = (a+4)^{n+1/2}\frac{1}{2\sqrt{\pi}n^{3/2}}.$$

Just note that $a = 0$ is the well-known formula for binary trees with n nodes.

Now we move to the generating function for the average number of registers. Apart from normalization it is

$$\sum_{p\geq 1}pR_p(z) = \sum_{p\geq 1}S_p(z) = \frac{1-u^2}{u}\sum_{p\geq 1}\sum_{k\geq 1}u^{k2^p}$$
$$= \frac{1-u^2}{u}\sum_{n\geq 1}v_2(n)u^n,$$

where $v_2(n)$ is the highest exponent k such 2^k divides n.

This has to be studied around $u = 1$, which, upon setting $u = e^{-t}$, means around $t = 0$. Eventually, and that is the only thing that is different here, this is to be retranslated into a singular expansion of z around its singularity, which depends on the parameter a.

For the reader's convenience, we also repeat the steps that were known before. The first factor is elementary:

$$\frac{1-u^2}{u} \sim 2t + \frac{1}{3}t^3 + \cdots$$

For

$$\sum_{p\geq 1}\sum_{k\geq 1} e^{-k2^p t},$$

one applies the Mellin transform, with the result

$$\frac{\Gamma(s)\zeta(s)}{2^s - 1}.$$

Applying the inversion formula, one finds

$$\sum_{p\geq 1}\sum_{k\geq 1} e^{-k2^p t} = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} t^{-s}\frac{\Gamma(s)\zeta(s)}{2^s - 1}ds.$$

Shifting the line of integration to the left, the residues at the poles $s = 1$, $s = 0$, $s = \chi_k = \frac{2k\pi i}{\log 2}$, $k \neq 0$ provide enough terms for our asymptotic expansion.

$$\frac{1}{t} + \frac{\gamma}{2\log 2} - \frac{1}{4} - \frac{\log \pi}{2\log 2} + \frac{\log t}{2\log 2} + \frac{1}{\log 2}\sum_{k\neq 0}\Gamma(\chi_k)\zeta(\chi_k)t^{-\chi_k}.$$

Combined with the elementary factor, this leads to

$$2 + \left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log t}{\log 2}\right)t + \frac{2t}{\log 2}\sum_{k\neq 0}\Gamma(\chi_k)\zeta(\chi_k)t^{-\chi_k} + O(t^2\log t).$$

Now we want to translate into the original z-world. Since $z = \frac{u}{1+(a+2)u+u^2}$, $u = 1$ translates into the singularity $z = \frac{1}{4+a}$. Further,

$$t \sim \sqrt{4+a}\cdot\sqrt{1-z(4+a)},$$

let us abbreviate $A = 4 + a$, then for singularity analysis we must consider

$$\frac{\sqrt{A}\cdot\sqrt{1-zA}\log(1-zA)}{2\log 2}$$

$$+ \left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log A}{2\log 2}\right)\sqrt{A}\cdot\sqrt{1-zA}$$

$$+ \frac{2}{\log 2}\sum_{k\neq 0}\Gamma(\chi_k)\zeta(\chi_k)A^{\frac{1-\chi_k}{2}}(1-zA)^{\frac{1-\chi_k}{2}}.$$

The formula that is perhaps less known and needed here is [5]

$$[z^n]\log(1-z)\sqrt{1-z} \sim \frac{n^{-3/2}\log n}{2\sqrt{\pi}} + \frac{n^{-3/2}}{2\sqrt{\pi}}(-2 + \gamma + 2\log 2);$$

furthermore we need

$$[z^n](1-z)^\alpha \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)}.$$

We start with the most complicated term:

$$\frac{[z^n]\frac{\sqrt{A}\cdot\sqrt{1-zA}\log(1-zA)}{2\log 2}}{[z^n]N(z)} \sim \frac{\sqrt{A}}{2\log 2}\frac{A^n\left(\frac{n^{-3/2}\log n}{2\sqrt{\pi}} + \frac{n^{-3/2}}{2\sqrt{\pi}}(-2+\gamma+2\log 2)\right)}{A^{n+1/2}\frac{1}{2\sqrt{\pi}n^{3/2}}}$$

$$= \log_4 n + 1 + \frac{\gamma}{2\log 2} - \frac{1}{\log 2}.$$

The next term we consider is

$$\left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log A}{2\log 2}\right)\sqrt{A}\frac{[z^n]\sqrt{1-zA}}{[z^n]N(z)}$$

$$\sim \left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log A}{2\log 2}\right)\sqrt{A}\frac{[z^n]\sqrt{1-zA}}{-\sqrt{A}[z^n]\sqrt{1-zA}}$$

$$= -\frac{\gamma}{\log 2} + \frac{1}{2} + \frac{\log \pi}{\log 2} - \frac{\log A}{2\log 2}.$$

The last term we consider is

$$\frac{2}{\log 2}\Gamma(\chi_k)\zeta(\chi_k)A^{\frac{1-\chi_k}{2}}\frac{[z^n](1-zA)^{\frac{1-\chi_k}{2}}}{-\sqrt{A}[z^n]\sqrt{1-zA}}$$

$$\sim -\frac{4\sqrt{\pi}}{\log 2}\frac{\Gamma(\chi_k)\zeta(\chi_k)}{\Gamma\left(\frac{\chi_k-1}{2}\right)}A^{\frac{1-\chi_k}{2}}n^{\chi_k/2}.$$

Eventually we have evaluated the average value of the Horton-Strahler numbers:

Theorem 1. *The average Horton-Strahler of weighted unary-binary trees with n nodes is given by the asymptotic formula*

$$\log_4 n - \frac{\gamma}{2\log 2} - \frac{1}{\log 2} + \frac{3}{2} + \frac{\log \pi}{\log 2} - \frac{\log A}{2\log 2} - \frac{4\sqrt{\pi A}}{\log 2}\sum_{k\neq 0}\frac{\Gamma(\chi_k)\zeta(\chi_k)}{\Gamma\left(\frac{\chi_k-1}{2}\right)}A^{\frac{-\chi_k}{2}}n^{\chi_k/2}$$

$$= \log_4 n - \frac{\gamma}{2\log 2} - \frac{1}{\log 2} + \frac{3}{2} + \frac{\log \pi}{\log 2} - \frac{\log A}{2\log 2} + \psi(\log_4 n),$$

with a tiny periodic function $\psi(x)$ of period 1.

4. MARKED ORDERED TREES

In [2] we find the following variation of ordered trees: Each rightmost edge might be marked or not, if it does not lead to an endnode (leaf). We depict a marked edge by the red colour and draw all of them of size 4 (4 nodes):
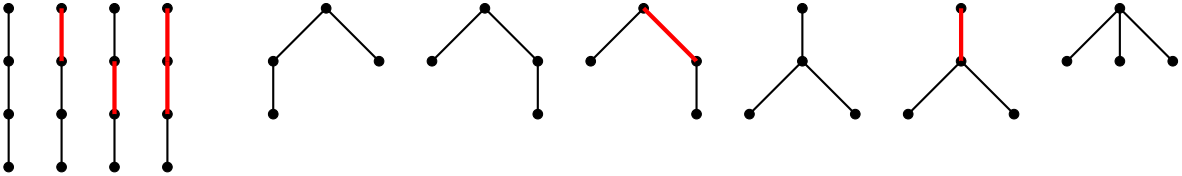
FIGURE 1. All 10 marked ordered trees with 4 nodes.

Now we move to a symbolic equation for the marked ordered trees:

FIGURE 2. Symbolic equation for marked ordered trees.
$\mathscr{A}\cdots\mathscr{A}$ refers to ≥ 0 copies of $\mathscr{A}$.

In terms of generating functions,

$$A = z + \frac{z}{1 - A}z + \frac{z}{1 - A}2(A - z),$$

with the solution

$$A(z) = \frac{1 - z - \sqrt{1 - 6z + 5z^2}}{2} = z + z^2 + z^3 + 3z^3 + 10z^4 + 36z^5 + \cdots.$$

The importance of this family of trees lies in the bijection to skew Dyck paths, as given in [2]. One walks around the tree as one usually does and translates it into a Dyck path. The only difference are the red edges. On the way down, nothing special is to be reported, but on the way up, it is translated into a skew step $(-1, -1)$. The present author believes that trees are more manageable when it comes to enumeration issues than skew Dyck paths.

The 10 trees of Figure 1 translate as follows:
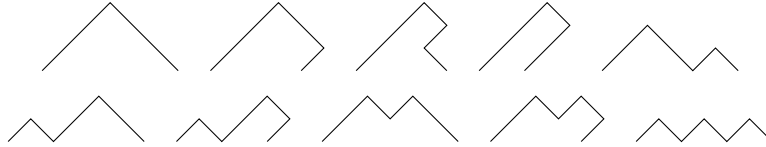
5. PARAMETERS OF MARKED ORDERED TREES

There are many parameters, usually considered in the context of ordered trees, that can be considered for marked ordered trees. Of course, we cannot be encyclopedic, also, to keep the balance between the other structures that are considered in this paper. We just consider a few parameters and leave further analysis to the future.

The number of leaves. To get this, it is most natural to use an additional variable u when translating the symbolic equation, so that $z^n u^k$ refers to trees with n nodes and k leaves. One obtains

$$F = zu + \frac{z}{1 - F}\big(zu2(F - zu)\big),$$

with the solution

$$F(z, u) = -z + \frac{zu}{2} + \frac{1}{2} - \frac{1}{2}\sqrt{4z^2 - 4z + z^2 u^2 - 2zu + 1}$$
$$= zu + z^2 u + \big(2u + u^2\big)z^3 + \big(4u + 5u^2 + u^3\big)z^4 + \cdots.$$

The factor $4u + 5u^2 + u^3$ corresponds to the 10 trees in Figure 1.

Of interest is also the average number of leaves, when all marked ordered trees of size n are considered to be equally likely. For that, we differentiate $F(z, u)$ w. r. t. u, followed by $u := 1$, with the result

$$\frac{z}{2} + \frac{z - z^2}{2\sqrt{1 - 6z + 5z^2}} = \frac{z}{1 - v}, \quad \text{with the usual} \quad z = \frac{v}{1 + 3v + v^2}.$$

Since $F(z, 1) = z(1 + v)$, it follows that the average is asymptotic to

$$\frac{[z^{n+1}]\frac{z}{1-v}}{[z^{n+1}]z(1 + v)} = \frac{[z^n]\frac{1}{1-v}}{[z^n](1 + v)} = \frac{[z^n]\frac{1}{\sqrt{5}}\frac{1}{\sqrt{1-5z}}}{5^{n+\frac{1}{2}}\frac{1}{2\sqrt{\pi}}n^{3/2}}$$

$$= \frac{\frac{n^{-1/2}}{\Gamma(\frac{1}{2})}}{5^{n+\frac{1}{2}}\frac{1}{2\sqrt{\pi}}n^{3/2}} = \frac{n}{10}.$$

Note that the corresponding number for ordered trees (unmarked) is $\frac{n}{2}$, so we have significantly less leaves here.

The height. As in the seminal paper [1], we define the height in terms of the longest chain of nodes from the root to a leaf. Further, let $p_h = p_h(z)$ be the generating function of marked ordered trees of height $\leq h$. From the symbolic equation,

$$p_{h+1} = z + \frac{z^2}{1 - p_h} + \frac{2z^2(p_h - z)}{1 - p_h} = -z + \frac{2z - z^2}{1 - p_h}, \quad h \geq 1, \ p_1 = z.$$

By some creative guessing, separating numerator and denominator, we find the solution

$$p_h = z(1+v)\frac{(1+2v)^{h-1} - v^h(v+2)^{h-1}}{(1+2v)^{h-1} - v^{h+1}(v+2)^{h-1}},$$

which is easy to prove by induction. The limit for $h \to \infty$ is $z(1+v)$, the generating function of *all* marked ordered trees, as expected. Taking differences, we get the generating functions of trees of height $> h$:

$$z(1+v) - z(1+v)\frac{(1+2v)^{h-1} - v^h(v+2)^{h-1}}{(1+2v)^{h-1} - v^{h+1}(v+2)^{h-1}}$$

$$= z(1+v)\frac{(1+2v)^{h-1} - v^{h+1}(v+2)^{h-1} - (1+2v)^{h-1} + v^h(v+2)^{h-1}}{(1+2v)^{h-1} - v^{h+1}(v+2)^{h-1}}$$

$$= z(1-v^2)\frac{(v+2)^{h-1}v^h}{(1+2v)^{h-1} - v^{h+1}(v+2)^{h-1}}.$$

From this, the average height can be worked out, as in the model paper [8]. We sketch the essential steps. For the average height, one needs

$$\sum_{h \geq 0} z(1-v^2)\frac{(v+2)^{h-1}v^h}{(1+2v)^{h-1} - v^{h+1}(v+2)^{h-1}}$$

and its behaviour around $v = 1$, viz.

$$2z(1-v)\sum_{h \geq 0}\frac{3^{h-1}v^h}{3^{h-1} - v^{h+1}3^{h-1}} \sim 2z(1-v)\sum_{h \geq 1}\frac{v^h}{1 - v^h}.$$

The behaviour of the series can be taken straight from [8].

We find there

$$\sum_{h \geq 1}\frac{v^h}{1 - v^h} = -\frac{\log(1 - v)}{1 - v},$$

and further

$$\sum_{h \geq 0} z(1-v^2)\frac{(v+2)^{h-1}v^h}{(1+2v)^{h-1} - v^{h+1}(v+2)^{h-1}} \sim -2z\log(1 - v),$$

so that the coefficient of z^{n+1} is asymptotic to $-2[z^n]\log(1 - v)$. Since $1 - v \sim \sqrt{5}\sqrt{1 - 5z}$,

$$-2z\log(1 - v) \sim -2z\log\sqrt{1 - 5z} = -z\log(1 - 5z),$$

and the coefficient of z^{n+1} in it is asymptotic to $\frac{5^n}{n}$. This has to be divided (as derived earlier) by

$$5^{n+\frac{1}{2}}\frac{1}{2\sqrt{\pi}n^{3/2}},$$

with the result

$$2\frac{5^n}{n}\frac{1}{5^{n+\frac{1}{2}}}\sqrt{\pi}n^{3/2} = \frac{2}{\sqrt{5}}\sqrt{\pi n}.$$

Note that the constant in front of $\sqrt{\pi n}$ for ordered trees is $\frac{2}{\sqrt{4}} = 1$, so the average height for marked ordered trees is indeed a bit smaller thanks to the extra markings.

6. A bijection between multi-edge trees and 3-coloured Motzkin paths

Multi-edge trees are like ordered (planar, plane, ...) trees, but instead of edges there are multiple edges. When drawing such a tree, instead of drawing, say 5 parallel edges, we just draw one edge and put the number 5 on it as a label. These trees were studied in [4, 8]. For the enumeration, one must count edges. The generating function $F(z)$ satisfies

$$F(z) = \sum_{k\geq 0}\left(\frac{z}{1-z}F(z)\right)^k = \frac{1}{1-\frac{z}{1-z}F(z)},$$

whence

$$F(z) = \frac{1-z-\sqrt{1-6z+5z^2}}{2z} = 1 + z + 3z^2 + 10z^3 + 36z^4 + 137z^5 + 543z^6 + \cdots.$$

The coefficients form once again sequence A002212 in [12].

A Motzkin path consists of up-steps, down-steps, and horizontal steps, see sequence A091965 in [12] and the references given there. As Dyck paths, they start at the origin and end, after n steps again at the x-axis, but are not allowed to go below the x-axis. A 3-coloured Motzkin path is built as a Motzkin path, but there are 3 different types of horizontal steps, which we call *red, green, blue*. The generating function $M(z)$ satisfies

$$M(z) = 1 + 3zM(z) + z^2M(z)^2 = \frac{1-3z-\sqrt{1-6z+5z^2}}{2z^2}, \quad \text{or} \quad F(z) = 1 + zM(z).$$

So multi-edge trees with N edges (counting the multiplicities) correspond to 3-coloured Motzkin paths of length $N-1$.

The purpose of this note is to describe a bijection. It transforms trees into paths, but all steps are reversible.

The details. As a first step, the multiplicities will be ignored, and the tree then has only n edges. The standard translation of such tree into the world of Dyck paths, which is in every book on combinatorics, leads to a Dyck path of length $2n$. Then the Dyck path will transformed bijectively to a 2-coloured Motzkin path of length $n-1$ (the colours used are red and green). This transformation plays a prominent role in [3], but is most likely much older. I believe that people like Viennot know this for 40 years. I would be glad to get a proper historic account from the gentle readers.

The last step is then to use the third colour (blue) to deal with the multiplicities.

The first up-step and the last down-step of the Dyck path will be deleted. Then, the remaining $2n - 2$ steps are coded pairwise into a 2-Motzkin path of length $n - 1$:

The last step is to deal with the multiplicities. If an edge is labelled with the number a, we will insert $a - 1$ horizontal blue steps in the following way: Since there are currently $n-1$ symbols in the path, we have n possible positions to enter something (in the beginning, in the end, between symbols). We go through the tree in pre-order, and enter the multiplicities one by one using the blue horizontal steps.

To make this procedure more clear, we prepared a list of 10 multi-edge trees with 3 edges, and the corresponding 3-Motzkin paths of length 2, with intermediate steps completely worked out:

Connecting unary-binary trees with multi-edge trees. This is not too difficult: We start from multi-edge trees, and ignore the multiplicities at the moment. Then we apply the classical rotation correspondence (also called: natural correspondence). Then we add vertical edges, if the multiplicity is higher than 1. To be precise, if there is a node, and an edge with multiplicity a leads to it from the top, we insert $a-1$ extra nodes in a chain on the top, and connect them with unary branches. The following example with 10 objects will help to understand this procedure. After that, all the structures studied in this paper are connected with bijections.

References

[1] N. G. De Bruijn, D. E. Knuth, and S. O. Rice. The average height of planted plane trees. In R. C. Read, editor, *Graph Theory and Computing*, pages 15–22. Academic Press, 1972.

[2] E. Deutsch, E. Munarini, S. Rinaldi. Skew Dyck paths. *J. Stat. Plann. Infer.*, 140 (8) (2010) 2191–2203.

[3] Emeric Deutsch and Louis W. Shapiro. A bijection between ordered trees and 2-Motzkin paths and its many consequences. volume 256, pages 655–670. 2002. LaCIM 2000 Conference on Combinatorics, Computer Science and Applications (Montreal, QC).

[4] M. Dziemiańczuk. Enumerations of plane trees with multiple edges and Raney lattice paths. *Discrete Math.*, 337:9–24, 2014.

[5] P. Flajolet and A. M. Odlyzko. Singularity analysis of generating functions. 3:216–240, 1990.

[6] P. Flajolet and H. Prodinger. Register allocation for unary-binary trees. *SIAM J. Comput.*, 15(3):629–640, 1986.

[7] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.

[8] Clemens Heuberger, Helmut Prodinger, and Stephan Wagner. The height of multiple edge plane trees. *Aequationes Math.*, 90(3):625–645, 2016.

[9] Hana Kim and Richard P. Stanley. A refined enumeration of hex trees and related polynomials. *European J. Combin.*, 54:207–219, 2016.

[10] H. Prodinger. Philippe Flajolet's early work in combinatorics. 2:1 (2022) Article S1H1.:11 pages, 2022.

[11] Helmut Prodinger. Introduction to Philippe Flajolet's work on the register function and related topics. *https://www.math.tugraz.at/∼prodinger/flajolet-introduction-register.pdf*, 7 pages, 2025.

[12] Neil J. A. Sloane and The OEIS Foundation Inc. The on-line encyclopedia of integer sequences, 2021.

Multi-edge tree	Dyck path	2-Motzkin path	blue edges added

TABLE 1. First row is a multi-edge tree with 3 edges, second row is the standard Dyck path (multiplicities ignored), third row is cutting off first and last step, and then translated pairs of steps, fourth row is inserting blued horizontal edges, according to multiplicities.

HELMUT PRODINGER, MATHEMATICS DEPARTMENT, STELLENBOSCH UNIVERSITY, 7602 STEL-LENBOSCH, SOUTH AFRICA.

Email address: `hproding@sun.ac.za`
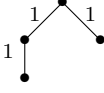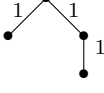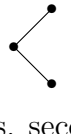
Multi-edge tree	Binary tree (rotation)	vertical edges added

TABLE 2. First row is a multi-edge tree with 3 edges, second row the corresponding binary tree, according to the classical rotation correspondence, ignoring the unary branches. Third row is inserting extra horizontal edges when the multiplicities are higher than 1.