ON THE NUMBER OF INCREASING TREES WITH LABEL REPETITIONS

OLIVIER BODINI, ANTOINE GENITRINI, AND BERNHARD GITTENBERGER

Abstract. In this paper we study a special subclass of monotonically labeled increasing trees: each sequence of labels from the root to any leaf is strictly increasing and each integer between 1 and k must appear in the tree, where k is the largest label. The main difference with the classical model of binary increasing tree is that the same label can appear in distinct branches of the tree. Such a class of trees can be used in order to model population evolution processes or concurrent processes.

A specificity of such trees is that they are built through an evolution process that induces ordinary generating functions. Finally, we solve the nice counting problem for these trees of size n and observe interesting asymptotics involving powers of n with irrational exponents.

1. Introduction

A rooted binary plane tree of size n is said *monotonically increasingly labeled* with the integers in $\{1, 2, \dots, k\}$ if each sequence of labels from the root to any leaf is weakly increasing. The concept of a monotonically labeled tree (of fixed arity $t \geq 2$) has been introduced in the 80's by Prodinger and Urbanek [11] and has then been revisited by Blieberger [1] in the context of Motzkin trees. In the latter paper, monotonically labeled Motzkin trees are directly related to the enumeration of expression trees that are built during compilation or in symbolic manipulation systems.

A rooted binary plane tree of size n is said *increasingly labeled* with the integers in $\{1, 2, \dots, n\}$ if each sequence of labels from the root to any leaf is increasing and each integer from 1 to n appears exactly once in the tree. Thus the sequences of labels along the branches are strictly increasing. This model corresponds to the heap data structure in computer science and is also related to the classical binary search tree model.

In this paper we are interested in a model lying between the two previous ones. It is in fact a special subclass of monotonically labeled increasing trees: *each sequence of labels from the root to any leaf is strictly increasing and each integer between* 1 *and* k *must appear in the tree*, where k is the largest label. The main difference from the classical model of binary increasing trees (*cf.* e.g. the book of Drmota [5]) is that the same label can appear in distinct branches of the tree. Our interest in such a model relies on the following fact: There is a classical *evolution process*, presented for example in [5], to grow a binary increasing tree replacing at each step an unlabeled leaf by an internal node labeled by the step number and attached to two new leaves. Here, we extend the process by selecting at each step a subset of leaves and replacing each of them by the same structure (the labeled internal nodes, all with the same

Date: September 13, 2018.

1991 *Mathematics Subject Classification.* Primary: 05A16, Secondary: 05C05, 34E05.

Key words and phrases. increasing trees, ordinary differential equations.

This work was partially supported by the ANR project *MetACOnc* ANR-15-CE40-0014, by the FWF grant SFB F50-03 and by the PHC Amadeus project 39454SF.

integer label, and their two children), thus an increasing binary tree with repetitions is under construction.

Such a model is totally natural to describe population evolution processes where each individual can give birth to two descendants independently of the other individuals. In another paper [4] we have presented an increasing model with repetitions of Schröder trees that encodes the chronology in phylogenetic trees.

Finally, by merging the nodes with the same label we obtain directed acyclic graphs whose nodes are increasingly labeled (without repetitions). Such an approach introduces a new model of concurrent processes with synchronization that induces processes whose description is more expressive than the classical series-parallel model that we have studied in [3, 2].

Outline of the paper. In Section 2 we introduce the concept of increasing binary trees with repetitions. In particular, we develop the evolution process naturally defining such trees. We also present the asymptotic behavior of their enumeration sequence. Section 3 is devoted to the asymptotic study of the number of increasing binary trees with repetitions of size n. Starting from the recurrence we first get the exponential growth rate by a purely heuristic approach. Afterwards, we use this guess to define an auxiliary sequence which is simpler to analyze. We derive a differential equation for the generating function the solutions of which are then asymptotically analyzed. Eventually, the asymptotic behavior of the auxiliary sequence can be determined. This leads then to the asymptotic behavior of the number of increasing binary trees with repetitions of size n and justifies and proves our guess after all. Then, in Section 4 we present a brief discussion of the generalization to k-ary trees.

2. Basic concepts and statement of the main result

The concept of an *increasing tree* is well studied in the literature (*cf.* for example [5]). An increasing tree is defined as a rooted labeled tree where on each path from the root to a leaf the sequence of labels is increasing. In fact they are strictly increasing, since the nodes of a labeled tree with n nodes carry exactly the labels $1, 2, \ldots, n$. The aim of the paper is to introduce a weaker model of increasing trees where repetitions of the labels can appear.

Definition 1. An increasing binary tree with repetitions is

- a binary tree, that is not necessarily complete, *i.e.*, the nodes have arity 0, 1 (with two possibilities: either a left child or a right one) or 2;
- the nodes are labeled according to the following constraints:
  - If a node has label k, then all integers from 1 to $k-1$ appear as labels in the tree. The set of labels is therefore a complete interval of integers of the form $\{1, 2, \ldots, m\}$ where m is the maximal label occurring in the tree.
  - Along each branch, starting from the root, the sequence of labels is (strictly) increasing.

We can complete an increasing binary tree with repetitions by plugging to each node whose arity is smaller than 2 either one or two leaves (without any label) to reach arity 2 to all the labeled nodes. We define the size of an increasing binary tree with repetitions as the number of leaves in the completed binary tree. This definition of the size will be completely natural once we will have introduced the way of constructing such trees.

In the rest of the paper, an increasing binary tree with repetitions will be called *weakly increasing tree*. In the Figure 1 a tree and its associated completed tree are represented.

Their common size is 8. If we want to expand the tree further, some of the •-leaves will take the label 5.

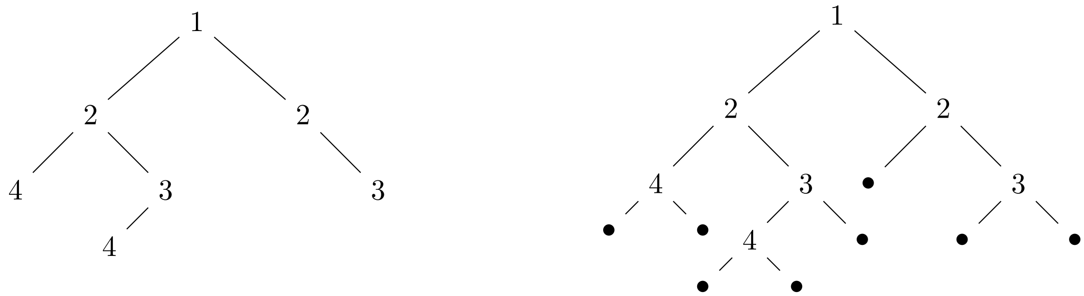

FIGURE 1. Left: a weakly increasing tree with 7 nodes and 4 distinct labels; Right: its associated completed tree.

Let us introduce a *combinatorial evolution process* to build weakly increasing trees. In order to construct a weakly increasing tree whose greatest label is m, start with the size-2 weakly increasing tree (a root with label 1) and repeat $(m-1)$ times the following step. At step $i \in \{1, \ldots, m-1\}$, choose a non-empty subset of •-leaves from the current completed weakly increasing tree and replace each of them by a tree with root labeled by $(i+1)$ and two •-leaves. At the end of the process remove the •-leaves to formally obtain a weakly increasing tree.

Lemma 2. *A weakly increasing tree being fixed, there is a single evolution that builds it.*

The combinatorial structure of weakly increasing trees being now formally defined, we denote by B_n the number of weakly increasing trees of size n, then we will prove the following quantitative result.

Theorem 3. *The number of weakly increasing binary trees of size n is asymptotically given by*

$$B_n \underset{n\to\infty}{\sim} \gamma\, n^{-\ln 2} \left(\frac{1}{\ln 2}\right)^n (n-1)!,$$

where γ is a constant.

This result can be compared to the number of classical increasing binary trees (without label repetition) with $n-1$ labeled nodes given by the number $(n-1)!$. For the latter model, the reader can refer to Flajolet and Segdewick's book [8, p. 143].

3. Enumeration of weakly increasing binary trees

Using the combinatorial evolution process to build trees (described in the previous section) and its associated Lemma 2, we get directly a recurrence for the partition of the size-n weakly increasing trees according to their maximal label m:

$$\begin{aligned} B_{1,2} &= 1 \\ B_{1,n} &= 0 \qquad \text{if } n \neq 2 \\ B_{m,n} &= \sum_{\ell=1}^{\lfloor (n-m)/2 \rfloor} \binom{n-\ell}{\ell} B_{m-1,n-\ell}, \end{aligned} \tag{1}$$

where $B_{m,n}$ is the number of weakly increasing trees with n nodes in which exactly m distinct labels occur. We remark that $B_n = \sum_{m\geq 1} B_{m,n}$, and thus the first terms of B_n are

$$0, 0, 1, 2, 7, 34, 214, 1652, 15121, 160110, 1925442, 25924260, 386354366, 6314171932, \dots$$

The first term of our sequence coincide with those of a shifted version of the sequence `A171792` in OEIS[1]. Some properties of this sequence are stated there, but no combinatorial meaning is given.

3.1. **The generating function of the counting sequence.** Let us now introduce the *ordinary generating series* $B(z)$ associated to the sequence B_n:

$$B(z) = \sum_{n\geq 0} B_n z^n.$$

The variable z marks the $\bullet$-leaves in the weakly increasing trees. Although the trees are labeled, we use an ordinary generating series because in this context, the evolution process directly turns to a combinatorial specification satisfied by the series $B(z)$. Recall, at each step, some $\bullet$-leaves are replaced by a deterministic labeled node with two $\bullet$-leaves, thus we get:

$$B(z) = z^2 + B(z + z^2) - B(z). \tag{2}$$

In fact, a tree is either the smallest tree (the root labeled by 1 with two $\bullet$-leaves) or it is obtained after the expansion of a tree where some $\bullet$-leaves do not change, and the other ones are replaced by a labeled node and two $\bullet$-leaves: $z \to z^2$). But at each step at least one leaf must be chosen, thus we remove the trees where no leaf has been chosen; these have generating function $B(z)$. The functional equation (2) can be rewritten as

$$B(z) = \frac{1}{2}\left(z^2 + B\left(z + z^2\right)\right). \tag{3}$$

Considering Equation (3), we prove that our sequence $(B_n)_{n\geq 0}$ is a shifted version of OEIS `A171792`.

It is remarkable that the most natural description here uses ordinary generating functions, unlike the exponential ones that are used for classical increasing trees. Furthermore, obtaining a functional equation satisfied by the exponential generating function associated to $(B_n)_n$ is not immediate. We may applying the combinatorial Borel transform on equation (3), which translates $B(z)$ into its exponential counterpart. But this does not seem to reveal another natural combinatorial way for defining weakly increasing trees. When deriving the asymptotics, however, we will start to work with the exponential generating function in order to deal with analytic functions.

3.2. **Analysis of the functional equation – heuristics.** Now we turn to the actual enumeration problem which amounts to the analysis of the function given in (3). First, we read off coefficients in (3) and get a recurrence relation for $B_n = [z^n]B(z)$, which is, of course, in

[1]OEIS means *Online Encyclopedia of Integer Sequences* that can be reached at `http://oeis.org/`

compliance with (1):

$$B_n = \sum_{\ell=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n-\ell}{\ell} B_{n-\ell} \tag{4}$$

$$= \sum_{p=n-\lfloor \frac{n}{2} \rfloor}^{n-1} \binom{p}{n-p} B_p. \tag{5}$$

Looking at this recurrence, we immediately observe that $B_n \geq (n-1)!$, thus $B(z)$ is only a formal power series. To get a first guess of the asymptotic behavior of B_n, we start with the following heuristic consideration: *Assume that* $B_n \underset{n\to\infty}{\sim} \alpha^n\, n!$*, for some* $\alpha > 0$.

Then the asymptotic analysis of B_n could be done by a singularity analysis (see in particular [7, 8]) of the exponential generating function

$$\hat{B}(z) = \sum_{n \geq 0} B_n \frac{z^n}{n!}.$$

Set $\phi_n := \alpha^n\, n!$ and add B_n to both sides of equation (5). This gives, when summing up over all n and assuming $B_n = \phi_n$, based on the left-hand side of (5)

$$2 \sum_{n \geq 0} B_n \frac{z^n}{n!} = 2 \sum_{n \geq 0} \phi_n \frac{z^n}{n!} = \frac{2}{1-\alpha z}.$$

By using the right-hand side of (5) we deduce

$$\frac{2}{1-\alpha z} = \sum_{n \geq 0} \frac{z^n}{n!} \sum_{p=n-\lfloor \frac{n}{2} \rfloor}^{n} \binom{p}{n-p} \phi_p$$

$$= \sum_{n \geq 0} z^n \sum_{p=n-\lfloor \frac{n}{2} \rfloor}^{n} \frac{\phi_p}{p!} \cdot \frac{1}{(n-p)!} \cdot \frac{p!^2}{n!(2p-n)!}.$$

Note that $\frac{p!^2}{n!(2p-n)!} \approx 1$ for $p \approx n$. If p is getting smaller then $\frac{p!^2}{n!(2p-n)!}$ rapidly tends to 0. And so does $\frac{1}{(n-p)!}$. Thus, only the last few terms of the inner sum should already almost give its value. Thus, let us assume that

$$\frac{2}{1-\alpha z} \sim \sum_{n \geq 0} z^n \sum_{p=0}^{n} \frac{\phi_p}{p!} \cdot \frac{1}{(n-p)!} = \frac{e^z}{1-\alpha z}.$$

But now, we see that both generating functions have a unique dominant singularity at $1/\alpha$ and they are approximately the same function. Thus, as $z \to 1/\alpha$, we must have that $2 \sim e^{1/\alpha}$, which yields $\alpha = 1/\ln 2 \approx 1.442695041\ldots$.

Note, that the reasoning above is only heuristic. There are many imprecisions in our arguments, so we have not proved anything so far. However, comparing $(\ln 2)^{-n}\, n!$ with the first 1000 values of (B_n) indicates that $B_n \sim b_n (\ln 2)^{-n} (n-1)!$ where $b_n \to 0$ at a slower rate than $1/n$.

3.3. **Analysis of the functional equation – asymptotics.** With the heuristic observation of the last section in mind, we define a new sequence $(b_n)_{n\geq 2}$ by

$$B_n = b_n \left(\frac{1}{\ln 2}\right)^n (n-1)!$$

and start to analyze it. According to our heuristic we set $\alpha = 1/\ln 2$. Then, the recurrence (4) becomes

$$b_n = \sum_{\ell=1}^{\lfloor \frac{n}{2} \rfloor} \frac{\alpha^{-\ell}}{\ell!} \cdot \frac{(n-\ell)(n-\ell-1)\cdots(n-2\ell+1)}{(n-1)(n-2)\cdots(n-\ell)} \cdot b_{n-\ell}. \tag{6}$$

To proceed, we need to analyze the second factor. We easily get a first bound.

Lemma 4. *The sequence $(b_n)_{n\geq 0}$ defined by $b_0 = b_1 = 0$, $b_2 = 1$, and (6) for $n \geq 3$ satisfies $b_n \leq 2/n^{1/\alpha}$.*

Proof. The estimate holds for $n \leq 2$. In particular, this implies $b_n \leq 1$ for $n \leq 2$. Using (6) and $\frac{(n-\ell)(n-\ell-1)\cdots(n-2\ell+1)}{(n-1)(n-2)\cdots(n-\ell)} \leq 1$ we obtain by induction

$$b_n \leq 2\sum_{\ell=1}^{n} \frac{\alpha^{-\ell}}{\ell!} \cdot \frac{1}{(n-\ell)^{1/\alpha}} = \frac{2}{n^{1/\alpha}} \sum_{\ell=1}^{n} \frac{\alpha^{-\ell}}{\ell!} \left(1 - \frac{\ell}{n}\right)^{-1/\alpha} \leq \frac{2}{n^{1/\alpha}} \left(e^{1/\alpha} - 1\right) = \frac{2}{n^{1/\alpha}}.$$

□

Our goal is to define two further sequences, one dominating b_n and the other dominated by b_n, but both having the same asymptotic behavior and hence b_n, too.

Lemma 5. *Set $\gamma_{n,\ell} := \frac{(n-\ell)(n-\ell-1)\cdots(n-2\ell+1)}{(n-1)(n-2)\cdots(n-\ell)}$. Then we have*

$$1 - \frac{\ell(\ell-1)}{n} - \frac{\ell(\ell-1)^2}{n^2} \leq \gamma_{n,\ell} \leq 1 - \frac{\ell(\ell-1)}{n} + \frac{\ell^2(\ell-1)(\ell-2)}{2n^2}.$$

Proof. To show the upper bound for $\gamma_{n,\ell}$, observe, in the definition of $\gamma_{n,\ell}$, the factor $n-\ell$ can be canceled in the numerator and in the denominator. Then write

$$\begin{aligned} \gamma_{n,\ell} &= \left(1 - \frac{\ell}{n-1}\right)\left(1 - \frac{\ell}{n-2}\right)\cdots\left(1 - \frac{\ell}{n-\ell+1}\right) \\ &\leq \left(1 - \frac{\ell}{n}\right)^{\ell-1} \end{aligned} \tag{7}$$

Using the fact that $(1-x)^n \leq 1 - nx + \binom{n}{2}x^2$, we get the stated upper bound.

Now we turn to the lower bound for $\gamma_{n,\ell}$. The sequence of harmonic numbers being denoted by $(H_n)_{n\geq 1}$, recall (*cf.* [12]) that the sequence $(H_n - \ln n)_n$ is monotonically decreasing. Thus, starting from (7) we obtain

$$\begin{aligned} \gamma_{n,\ell} &\geq 1 - \ell\left(\frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-\ell+1}\right) = 1 - \ell\left(H_{n-1} - H_{n-\ell}\right) \\ &\geq 1 - \ell\left(\ln(n-1) - \ln(n-\ell)\right) = 1 - \ell \ln\left(\frac{n-1}{n-\ell}\right) \\ &= 1 - \ell \ln\left(\frac{1}{1 - \frac{\ell-1}{n}}\right). \end{aligned} \tag{8}$$

Since $\gamma_{n,\ell}$ is defined for $\ell \le \lfloor n/2 \rfloor$, the result now follows from the inequality $\ln \frac{1}{1-x} \le x + x^2$, which holds for $0 \le x \le 1/2$. □

Next let us define the sequence $\overline{b}_n$ which dominates b_n.

Lemma 6. *Set* $\overline{b}_0 = \overline{b}_1 = 0$, $\overline{b}_2 = 1$, *and for all* $n \ge 3$,

$$\overline{b}_n = \sum_{\ell=1}^{n} \frac{\alpha^{-\ell}}{\ell!} \cdot \left(1 - \frac{\ell(\ell-1)}{n} + \frac{\ell^2(\ell-1)(\ell-2)}{2n^2}\right) \cdot \overline{b}_{n-\ell}, \tag{9}$$

where $\alpha = 1/\ln 2$. *Then, we have* $\overline{b}_n \ge b_n$ *for all* $n \ge 0$.

Proof. Since the initial values for $\overline{b}_n$ and b_n are the same and in the recurrence $\gamma_{n,\ell}$ has been replaced by a larger expression as well as the range of summation is larger, the sequence $\overline{b}_n$ is obviously dominating the sequence b_n. □

We then define a sequence dominated by b_n.

Lemma 7. *Set* $\underline{b}_0 = \underline{b}_1 = 0$, $\underline{b}_2 = 1$, *and for all* $n \ge 3$,

$$\underline{b}_n = \sum_{\ell=1}^{n} \frac{\alpha^{-\ell}}{\ell!} \cdot \left(1 - \frac{\ell(\ell-1)}{n} - \frac{\ell(\ell-1)^2}{n^2}\right) \cdot \underline{b}_{n-\ell}, \tag{10}$$

where $\alpha = 1/\ln 2$. *Then, we have* $\underline{b}_n \le b_n$ *for all* $n \ge 0$.

Proof. When turning from b_n to $\underline{b}_n$, the initial values were not changed, but $\gamma_{n,\ell}$ was replaced by a smaller term. If the summation went from 1 to $\lfloor \frac{n}{2} \rfloor$, the new sequence would be obviously smaller than b_n. But note that the additional terms in the sum are negative, thus we indeed have $\underline{b}_n \le b_n$. □

The next step is deriving a differential equation satisfied by the generating function $\overline{B}(z) = \sum_{n\ge0} \overline{b}_n z^n$. Adding $\overline{b}_n$ on both sides of recurrence (9) implies

$$2\overline{B}(z) = \sum_{n\ge0} z^n \sum_{\ell=0}^{n} \overline{b}_{n-\ell} \frac{\alpha^{-\ell}}{\ell!} - \sum_{n\ge1} \frac{z^n}{n} \left(\sum_{\ell=0}^{n} \overline{b}_{n-\ell} \frac{\ell(\ell-1)\alpha^{-\ell}}{\ell!} + O\left(\frac{\alpha^{-\ell}}{\ell!} \cdot \frac{\ell^4}{n^2} \overline{b}_{n-\ell} \right) \right). \tag{11}$$

Now observe that

$$\sum_{\ell\ge0} \ell(\ell-1)\alpha^{-\ell} \frac{z^\ell}{\ell!} = \frac{z^2}{\alpha^2} e^{z/\alpha}.$$

Thus the Cauchy product on the right-hand side of (11) can be written in closed form and with the help of Lemma 4 we can estimate the error term and obtain

$$2\overline{B}(z) = e^{z/\alpha}\overline{B}(z) - \int_0^z \frac{x}{\alpha^2} e^{x/\alpha} \overline{B}(x)\, \mathrm{d}x + o\left(\ln \frac{1}{1-z} \right). \tag{12}$$

From this integral equation we can derive a first order differential equation which can be analyzed. Being more precise and replacing the error term by its exact value, we obtain the following result.

Lemma 8. *The generating function* $\overline{B}(z) = \sum_{n\ge0} \overline{b}_n z^n$ *satisfies the differential equation*

$$z\left(e^{z/\alpha} - 2\right)\overline{B}''(z) + \left(\left(1 - \frac{z}{\alpha}\right)^2 e^{z/\alpha} - 2\right)\overline{B}'(z) + \left(\frac{z^3}{2\alpha^4} + \frac{z^2}{2\alpha^3} - \frac{z}{\alpha^2} + \frac{1}{\alpha}\right) e^{z/\alpha}\overline{B}(z) = 0, \tag{13}$$

where $\alpha = 1/\ln 2$. Thus $\overline{B}(z)$ has a unique dominant singularity at $z = 1$ and exhibits the following asymptotic behavior as z tends to 1. There is a positive constant γ such that

$$\overline{B}(z) \underset{z\to 1}{\sim} \gamma\,(1-z)^{\frac{1}{\alpha}-1}. \tag{14}$$

Moreover, the coefficients of $\overline{B}(z)$ are asymptotically given by

$$\bar{b}_n = [z^n]\overline{B}(z) \underset{n\to\infty}{\sim} \gamma\,\frac{n^{-\frac{1}{\alpha}}}{\Gamma\left(1-\frac{1}{\alpha}\right)}. \tag{15}$$

Proof. Let us first consider the simpler problem (12) without considering the error term. A differentiation yields the functional-differential equation

$$\left(e^{z/\alpha}-2\right)\overline{B}'(z) - \left(\frac{z}{\alpha^2}-\frac{1}{\alpha}\right)e^{z/\alpha}\overline{B}(z) = 0. \tag{16}$$

This equation has the explicit generic solution

$$f(z) = C\left(e^{z/\alpha}-2\right)^{\frac{1}{\alpha}-1}\exp\left(-\int_1^{e^{z/\alpha}/2}\frac{\ln t}{1-t}\,\mathrm{d}t\right).$$

The function $f(z)$ is analytic in $|z|<1$ and its sole singularity on the circle of convergence is $z=1$. Since

$$e^{z/\alpha}-2 \underset{z\to 1}{\sim} -\frac{2}{\alpha}(1-z) + O(|1-z|^2),$$

and the second factor tends to 1, as z tends to 1, it indeed behaves like described in (14) and thus a standard transfer lemma [7, 8] yields the asymptotic behavior of its coefficients.

The error term in (11) is

$$\frac{1}{2}\sum_{n\geq 1}\frac{z^n}{n^2}\sum_{\ell=0}^{n}\bar{b}_{n-\ell}\frac{\ell^2(\ell-1)(\ell-2)\alpha^{-\ell}}{\ell!}.$$

Using the identity $\sum_\ell \ell^2(\ell-1)(\ell-2)\frac{z^\ell}{\ell!} = z^3(ze^z)''' = z^3(3+z)e^z$ we obtain the integral equation

$$z(e^{z/\alpha}-2)\overline{B}(z) - z\int_0^z \frac{x}{\alpha^2}e^{x/\alpha}\overline{B}(x)\,\mathrm{d}x + \frac{1}{2}\int_0^z\frac{1}{x}\int_0^x\frac{u^2}{\alpha^3}\left(3+\frac{u}{\alpha}\right)e^{u/\alpha}\overline{B}(u)\,\mathrm{d}u\,\mathrm{d}x,$$

which can be transformed into (13) by differentiating twice. This differential equation cannot be solved explicitly, but the coefficients of $\overline{B}(z)$, $\overline{B}'(z)$, and $\overline{B}''(z)$ are analytic functions. Thus, we can apply the standard theory of coefficient asymptotics of solutions of linear differential equations detailed in [8, Section VII. 9.1] and applied directly on ordinary differential equations, as laid out in the classical books of Ince [9] and Wasow [13]. To find the singularities of the solutions and the asymptotic behavior of the solutions near their singularities, we first have to look at the zeros of the highest order coefficient of the differential equation. This is $z\left(e^{z/\alpha}-2\right)$ and its zeros, $z=0$ and $z=1$, are not zeros of the other coefficients. Moreover, we observe that the difference of the two zeros is an integer.

The theory then tells us that there is a solution with a singularity at 1 and one with a singularity at 0. Due to the integer difference of the two singularities, the asymptotic expansions of these two solutions at their respective singular points may contain logarithmic factors. Since the asymptotic main terms must cancel, we easily find the corresponding exponents and possibly present logarithmic terms. One solution behaves like described in (14) when z is tending to 1, the other solution is asymptotically equivalent to $\frac{1}{z}\ln z$ when z is tending to 0. Since these two solutions span the whole space of solutions and the second one is obviously not a power series around 0, we obtain (14) and (15). □

For the lower bound we can do a similar reasoning.

Lemma 9. *The generating function* $\underline{B}(z) = \sum_{n\geq 0} \underline{b}_n z^n$ *has a unique dominant singularity at* $z = 1$ *and exhibits the following asymptotic behavior as* z *tends to* 1*. There is a positive constant* γ *such that*

$$\underline{B}(z) \underset{z\to 1}{\sim} \gamma\,(1-z)^{\frac{1}{\alpha}-1}.$$

Consequently, we have

$$\underline{b}_n = [z^n]\underline{B}(z) \underset{n\to\infty}{\sim} \gamma\,\frac{n^{-\frac{1}{\alpha}}}{\Gamma\left(1-\frac{1}{\alpha}\right)}.$$

Proof. We start with the lower bound for $\gamma_{n,\ell}$ given in Lemma 5. The reasoning in the proof of the previous lemma yields a second order linear differential equation. In the same way as before, it turns out that the terms of order $1/n^2$ only affect the coefficient of $\underline{B}(z)$ in the differential equation. Thus we know that $\underline{B}(z)$ satisfies the differential equation (16), up to a small perturbation not influencing the singular structure of the solutions. Hence $\underline{B}(z)$ and $\overline{B}(z)$ have the same dominant singularity and the same local behavior at their dominant singularity. Consequently, their coefficients are asymptotically equivalent. □

Thus, the asymptotic behavior of B_n stated in Theorem 3 is proved.

4. Higher arity weakly increasing trees

In this section we briefly discuss how our results generalize to k-ary weakly increasing trees with repetitions. The definitions for the binary case can be adapted in an obvious way. Then, we define the size to be the number of leaves of the completed k-ary tree and the generating function $B(z) = \sum_{n\geq k} B_n z^n$ where B_n is the number of k-ary weakly increasing trees (with repetitions) of size n. In a similar way as in the binary case we obtain the functional equation

$$B(z) = \frac{1}{2}\left(z^k + B(z+z^k)\right).$$

The first task is to get the exponential growth rate. Recall the recurrence for the binary case, $B_n = \sum_{\ell=1}^{\lfloor \frac{n}{2}\rfloor}\binom{n-\ell}{\ell}B_{n-\ell}$. This was obtained by expanding

$$B_n = \sum_{\ell=1}^{\lfloor \frac{n}{2}\rfloor} B_{n-\ell}[z^n](z+z^2)^{n-\ell} = \sum_{\ell=1}^{\lfloor \frac{n}{2}\rfloor} B_{n-\ell}[z^\ell](1+z)^{n-\ell}. \tag{17}$$

All we have to do now is replacing the z^2 on the right-hand side of the first line by z^k. From this, we get the recurrence for the k-ary case:

$$\begin{aligned}
B_n &= \sum_{\ell=1}^{\lfloor n-\frac{n}{k}\rfloor} B_{n-\ell}[z^n](z+z^k)^{n-\ell} = \sum_{\ell=1}^{\lfloor n-\frac{n}{k}\rfloor} B_{n-\ell}[z^\ell](1+z^{k-1})^{n-\ell} \\
&= \sum_{\substack{\ell=1,\\ \ell\equiv 0 \mod k-1}}^{\lfloor n-\frac{n}{k}\rfloor} \binom{n-\ell}{\frac{\ell}{k-1}} B_{n-\ell} = \sum_{s=1}^{\lfloor \frac{n}{k}\rfloor} B_{n-(k-1)s}\binom{n-(k-1)s}{s} \qquad (18)\\
&= \sum_{\substack{p=n-\lfloor \frac{n}{k}\rfloor,\\ p\equiv n \mod k-1}}^{n-1} \binom{p}{\frac{n-p}{k-1}} B_p.
\end{aligned}$$

The expressions in the last two lines show that only particular terms of the sequence $(B_n)_{n\ge 0}$ are used. This is no surprise, as the arity constraint implies that $B_n \neq 0$ if and only if $n \equiv 1 \mod k-1$. Thus we set $C_n = B_{1+n(k-1)}$. From (18) we obtain

$$C_n = \sum_{s=1}^{\lfloor n-\frac{n-1}{k}\rfloor} \binom{1+(n-s)(k-1)}{s} C_{n-s} = \sum_{s=\lceil \frac{n-1}{k}\rceil}^{n-1} \binom{1+s(k-1)}{n-s} C_s. \qquad (19)$$

Let us another time start with the heuristic approach. Take the first of these two equations for C_n, add C_n on both sides of the equation and assume that $C_n = \alpha^n n!$. Then, transforming both sides of the equation into their exponential generating functions gives

$$C(z) = \sum_{n\ge 0} C_n \frac{z^n}{n!} = \frac{2}{1-\alpha z}$$

for the left-hand side and

$$C(z) = \sum_{n\ge 0} \sum_{s=\lceil \frac{n-1}{k}\rceil}^{n} \binom{1+s(k-1)}{n-s} C_s \frac{z^n}{n!} = \sum_{n\ge 0} \sum_{s=\lceil \frac{n-1}{k}\rceil}^{n} \frac{C_s}{s!}\frac{1}{(n-s)!}\cdot \frac{(1+s(k-1))!s!}{(1+sk-n)!n!}$$

for the right-hand side. Reasoning as in the binary case (extending the inner sum to $0 \ge s \ge n$ and assuming that the last factor can be replaced by 1) gives $C(z) \sim e^z/(1-\alpha z)$ and $\alpha = 1/\ln 2$.
So, let us use the ansatz $C_n = c_n(\ln 2)^{-n} n!$. Then, by (19) we have

$$c_n = \sum_{s=1}^{\lfloor n-\frac{n-1}{k}\rfloor} \frac{(\ln 2)^s}{s!}\, \delta_{n,s}\, c_{n,s},$$

where

$$\begin{aligned}
\delta_{n,s} &= \frac{(1+(n-s)(k-1))!\,(n-s)!}{(1+(n-s)(k-1)-s)!\,n!}\\
&= \frac{((n-s)(k-1)+1)\cdot((n-s)(k-1))\cdots((n-s)(k-1)-s+2)}{n(n-1)\cdots(n-s+1)}. \qquad (20)
\end{aligned}$$

As in Lemma 4, a simple induction shows the bound $c_n \le 2(k-1)^n/n^{(k-1)(\ln 2-1)-1}$. In order to proceed, we have to get better estimates for $\delta_{n,s}$.

Lemma 10. *Let $\delta_{n,s}$ be defined by (20). Then we have*

$$(k-1)^s\left(1-\frac{s(s-1)}{n}-\frac{s(s-1)^2}{n^2}\right)\le \delta_{n,s}\le (k-1)^s\left(1-\frac{s(s-1)}{n}+\frac{s^2(s-1)(s-2)}{2n^2}\right).$$

Proof. All we have to show is that $\delta_{n,s}/(k-1)^s$ can be estimated similarly as $\gamma_{n,\ell}$ was estimated in (7) and (8). Indeed, since $k\ge 3$, we obtain

(21)

$$\begin{aligned}\frac{\delta_{n,s}}{(k-1)^s} &= \left(1-\frac{s}{n}+\frac{1}{n(k-1)}\right)\left(1-\frac{s-1}{n-1}\right)\cdots\left(1-\frac{s-(s-1)}{n-s+1}-\frac{s-2}{(n-s+1)(k-1)}\right)\\ &\le \left(1-\frac{s-1}{n}\right)\left(1-\frac{s-1}{n-1}\right)\cdots\left(1-\frac{s-1}{n-s+1}\right)\\ &\le \left(1-\frac{s-1}{n}\right)^s\end{aligned}$$

which is similar to (7) and implies the upper bound.

For the lower bound note that

$$\frac{s-\ell}{n-\ell}+\frac{\ell-1}{(n-\ell)(k-1)}\le\frac{s-1}{n-\ell}.$$

Inserting this into (21) and reasoning as after (8) we obtain

$$\begin{aligned}\frac{\delta_{n,s}}{(k-1)^s}&\ge 1-(s-1)\left(\frac{1}{n}-\frac{1}{n-1}-\cdots-\frac{1}{n-s+1}\right)\\ &\ge 1-(s-1)\ln\left(\frac{1}{1-\frac{s}{n}}\right)\end{aligned}$$

which gives the lower bound after all. □

As in the binary case, these two bounds can be used to define two sequences, one majorizing c_n and the other one minorizing c_n, which eventually show that $c(z)=\sum_{n\ge 0}c_nz^n$ satisfies

$$2c(z)\sim\sum_{n\ge 0}z^n\sum_{s=0}^{n}\frac{(\ln 2)^s(k-1)^s}{s!}c_{n-s}-\sum_{n\ge 1}\frac{z^n}{n}\sum_{s=0}^{n}\frac{(\ln 2)^s(k-1)^s}{s!}s(s-1)c_{n-s}. \tag{22}$$

which leads to the differential equation

$$\left(e^{z(k-1)\ln 2}-2\right)c'(z)-\left(z(k-1)^2(\ln 2)^2-(k-1)\ln 2\right)e^{z(k-1)\ln 2}c(z)=0.$$

As in the proof of Lemma 8, it turns out that the error term in (22) (which can be quantified by Lemma 10) yields a perturbation of the differential equation which does not affect the asymptotic behavior of the solutions near their dominant singularity. The only difference to the binary case is the factor $(k-1)$ popping up at several places. In fact, if we write the differential equation for $B(z)$ in the form $a(z)y'(z)+b(z)y(z)=0$, then the differential equation for $c(z)$ is $a(z(k-1))c'(z)+(k-1)b(z(k-1))c(z)=0$. Thus, the explicit solution of the differential equation for $B(z)$ yields eventually

$$c(z)\sim f((k-1)z)^{k-1},$$

where

$$f(z) = K\left(e^{z\ln 2} - 2\right)^{\ln 2 - 1} \exp\left(-\int_1^{e^{z\ln 2}/2} \frac{\ln t}{1-t}\mathrm{d}t\right)$$

for some constant K. Thus, we get the asymptotic behavior

$$C(z) \sim K\left(e^{z(k-1)\ln 2} - 2\right)^{(k-1)(\ln 2 - 1)}, \qquad \text{as } z \to \frac{1}{k-1},$$

which implies

$$c_n \underset{n\to\infty}{\sim} K(k-1)^n \frac{n^{-(k-1)\ln 2 + k - 2}}{\Gamma\left(-(k-1)(\ln 2 - 1)\right)}.$$

So, we get the following result after all.

Theorem 11. *The number of k-ary weakly increasing trees with repetitions which have size n is asymptotically given by*

$$B_n \begin{cases} = 0 & \text{if } n \not\equiv 1 \mod k-1, \\ \underset{n\to\infty}{\sim} \gamma\, m^{-(k-1)(\ln 2 - 1)} \left(\frac{k-1}{\ln 2}\right)^m (m-1)! & \text{if } n = 1 + (k-1)m, \end{cases}$$

where γ is a positive constant.

Remark that the latter theorem is coherent with the binary case after simplifications.

5. Conclusion

The problems in the context of classical increasing trees can often be translated into the context of urn models. Let us refer, for example, to the two following papers, among many ones, for the proximity with our study. The first by Mahmoud [10], relating urns and trees, and the second by Flajolet *et al.* [6] studying urns with Analytic Combinatorics.

There is a way to encode our binary case problem as an urn model with a single color for balls in the following way. *Start with an urn with two balls. At each step, sample a subset of r of balls in the urns, and then return $2r$ balls in the urn. How may histories are there that give an urn containing n balls?*

Unfortunately, to the best of our knowledge, urn processes with a random quantity of sampled balls have not been studied yet. Thus it seems very promising to us to develop this new model further.

References

[1] J. Blieberger. Monotonically labelled Motzkin trees. *Discrete Applied Mathematics*, 18(1):9–24, 1987.

[2] O. Bodini, M. Dien, A. Genitrini, and F. Peschanski. Entropic uniform sampling of linear extensions in series-parallel posets. In *12th International Computer Science Symposium in Russia (CSR)*, pages 71–84, 2017.

[3] O. Bodini, M. Dien, A. Genitrini, and F. Peschanski. The ordered and colored products in analytic combinatorics: application to the quantitative study of synchronizations in concurrent processes. In *14th SIAM Meeting on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 16–30, 2017.

[4] O. Bodini, A. Genitrini, and M. Naima. Ranked Schröder trees, 2018. Technical report.

[5] M. Drmota. *Random trees.* Springer, Vienna-New York, 2009.

[6] P. Flajolet, J. Gabarró, and H. Pekari. Analytic urns. *The Annals of Probability*, 33(3):1200–1233, 2005.

[7] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.

[8] P. Flajolet and R. Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[9] E. L. Ince. *Ordinary differential equations.* Dover Publications, New York, 1944.

[10] H. M. Mahmoud. Urn models and connections to random trees: A review. *Journal of the Iranian Statistical Society*, 2:53–114, 2003.
[11] H. Prodinger and F. J. Urbanek. On monotone functions of tree structures. *Discrete Applied Mathematics*, 5(2):223–239, 1983.
[12] S. R. Tims and J. A. Tyrrell. Approximate evaluation of Euler's constant. *The Mathematical Gazette*, 55(391):65–67, 1971.
[13] W. Wasow. *Asymptotic expansions for ordinary differential equations*. Dover Publications, Inc., New York, 1987. Reprint of the 1976 edition.

Laboratoire d'Informatique de Paris-Nord, CNRS UMR 7030 - Institut Galilée - Université Paris-Nord, 99, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France.
E-mail address: `Olivier.Bodini@lipn.univ-paris13.fr`

Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6 -LIP6- UMR 7606, F-75005 Paris, France.
E-mail address: `Antoine.Genitrini@lip6.fr`

Department of Discrete Mathematics and Geometry, Technische Universität Wien, Wiedner Hauptstrasse 8-10/104, 1040 Wien, Austria.
E-mail address: `gittenberger@dmg.tuwien.ac.at`