Entropy of Some General Plane Trees

Zbigniew Gołębiewski
Faculty of Fundamental Problems of Technology
Wroclaw University of Science and Technology
Wrocław, Poland
Email: zbigniew.golebiewski@pwr.edu.pl

Abram Magner
Coordinated Science Laboratory
Univ. of Illinois at Urbana-Champaign
Urbana, IL, USA
Email: anmagner@illinois.edu

Wojciech Szpankowski
Department of Computer Science
Purdue University
West Lafayette, IN, USA
Email: spa@cs.purdue.edu

***Abstract*—We continue developing information theory of advanced data structures. In our previous work, we introduced structural entropy of unlabeled graphs and designed lossless compression for *binary* trees (with correlated names). In this paper, we consider d-ary trees ($d \geq 2$) and trees with unrestricted degree for which we compute entropy (the first step to design optimal compression). As it turns out extending from binary trees to general trees is mathematically quite challenging and leads to new recurrences that find ample of applications in information theory (e.g., to analyze non-plane general trees).**

I. Introduction

Rapid advances in sensing, communication, and storage technologies have created a state of the art in which our ability to collect data from richly instrumented environments has far outpaced our ability to process, understand, and analyze this data in a (provably) rigorous manner. A significant component of this complexity arises from the multimodal and heterogeneous nature of data. This poses significant challenges for theoretical characterization of limits of information and methods that achieve these limits. While ad-hoc approaches are often currently deployed, critical issues regarding their performance, robustness, and scalability, remain. These precise challenges have motivated our recent research program [6], [7], [13] and others [2], [11], [18]). It provides the basis for our effort on developing a comprehensive theory of information for multimodal data, that is, multitype and context dependent structures.

As a start to understand advanced multimodal data structures, we focused on graphs [6] and trees [13]. In 1990, Naor proposed an efficiently computable representation for unlabeled graphs (solving Turán's open question) that is optimal up to the first two leading terms of the entropy when all unlabeled graphs are equally likely. Naor's result is asymptotically a special case of recent work of Choi and Szpankowski [6], who extended Naor's result to general Erdős-Rényi graphs. In particular, in [6] the entropy and an optimal compression algorithm (up to two leading terms of the entropy) for Erdős-Rényi graph structures were presented. Furthermore, in [14] an automata approach was used to design an optimal graph compression scheme. There also have been some heuristic methods for real-world graphs compression including grammar-based compression for some data structures. Peshkin [15] proposed an algorithm for a graphical extension of the one-dimensional SEQUITUR compression method. For binary plane-oriented trees rigorous information-theoretic results were obtained in [11], complemented by a universal grammar-based lossless coding scheme [18].

In our recent work [13] (see also [7]) we study *binary* trees (with correlated labels) and design an optimal compression based on arithmetic encoding. In this paper, we extend our study on entropy of advanced data structures to d-ary trees (i.e., trees with degree $d \geq 2$) and general trees without any restriction on degree. It turns out that moving from binary trees to d-ary (general) trees is mathematically quite challenging. First of all, in [13] we proved for binary trees an equivalence between two models: *binary search model* and a model in which leaves are selected randomly to expand the tree by adding two additional nodes (new leaves). These equivalence allowed us to analyze the entropy of such tree by solving a relatively simple recurrence, namely

$$x_n = a_n + \frac{2}{n}\sum_{i=1}^{n-1} x_i. \tag{1}$$

for some given a_n (e.g., for the entropy $a_n = \log n$), where n denotes the number of internal nodes. However, for d-ary trees T_n on n internal nodes the entropy $H(T_n)$ satisfies

$$H(T_n) = H(\text{root}) + d\sum_{k=0}^{n-1} H(T_k)p_{n,k}$$

where $H(\text{root})$ is the entropy of the split probability at the root, and $p_{n,k}$ is the probability of one specified subtree being of size k. This recurrence is quite simple for m-ary search tree model discussed in Section II, in which we store a permutation of $\{1, \ldots, n\}$ by splitting the file in the root using the first $m-1$ elements of the underlying permutation. In this case, $p_{n,k} = \binom{n-k-1}{m-2}/\binom{n}{m-1}$ and the recurrence was already discussed in [5], [9].

For more interesting d-ary trees we select randomly a leaf and add exactly d leaves to it. Observe that for $d = 2$ we

have $p_{n,k} = 1/n$ (where we recall n here denotes the number of internal nodes) leading to (1). But things are getting more complicated when $d > 2$. For example, for $d = 3$ we can prove that

$$p_{n,k} = \frac{1}{2n} \frac{\binom{2k}{k} 2^{2n}}{\binom{2n}{n} 2^{2k}}.$$

After some tedious algebra, we prove in Section III that the new type of recurrence we need to solve to find the entropy is of the following form (see Lemma 2)

$$x_n = a_n + \frac{\alpha}{n} \frac{n!}{\Gamma(n+\alpha-1)} \sum_{k=0}^{n-1} \frac{\Gamma(k+\alpha-1)}{k!} x_k \qquad (2)$$

where $\alpha = d/(d-1)$, a_n is given sequence, and Γ is the Euler gamma function. A situation is even more involved when we consider general trees in Section III-C where no restrictions on tree degree is imposed.

In this paper, after describing in Section II three possible generalizations of the binary tree model from [13], we present our main results in Section III. We first provide in Corollary 1 the entropy rate for m-ary search trees. Then we consider d-ary trees and in Theorem 2 give our expression for the entropy of such trees. We extend it to general trees in Theorem 3.

II. Models

In this section we describe the concepts of unlabeled plane trees with and without restrictions on the nodes out-degree. This will allow us to introduce three models of tree generations.

We call a rooted tree a plane tree when we distinguish left-to-right order of the successors of each node, i.e. we distinguish all different embedding of a tree into the plane (see [8]). Unlabeled plane trees are rooted plane trees with no restriction on the number of successors of the nodes. On the other hand, unlabeled d-ary plane trees are rooted plane trees where each node has exactly d successors (either internal of external). Since they are unlabeled, they can be seen as objects that encodes the structures of a possible plane tree. We define the size of the unlabeled plane tree and also unlabeled d-ary plane tree by the number of internal nodes.

A. Unlabeled m-Ary Search Trees Generation

Search trees are plane trees build from a set of n distinct keys taken from some totally ordered sets, for instance a random permutation of the numbers $\{1, 2, \ldots, n\}$. The search tree is m-ary tree where each node has at most m successors; moreover each node stores up to $m-1$ key in each node. We define the size of search tree as the number of keys n. Construction of m-ary search tree can be described as follows. If $n = 0$ the tree is empty. If $1 \leq n \leq m-1$ the tree consists of a root only, with all keys stored in the root. If $n \geq m$ we select $m-1$ keys that are called pivots. The pivots are stored in the root. The $m-1$ pivots split the set of remaining $n-m+1$ keys into m sublists $I_1, \ldots, I_m$: if the pivots are $p_1 < p_2 < \cdots < p_{m-1}$, then $I_1 := (p_i : p_i < p_1), I_2 := (p_i : p_1 < p_i < p_2), \ldots, I_m := (p_i : p_{m-1} < p_i)$. We then recursively construct a search tree for each of the sets I_i of keys. In order to obtain unlabeled search tree of size n we remove keys from the search trees of size n (see Fig. 1).

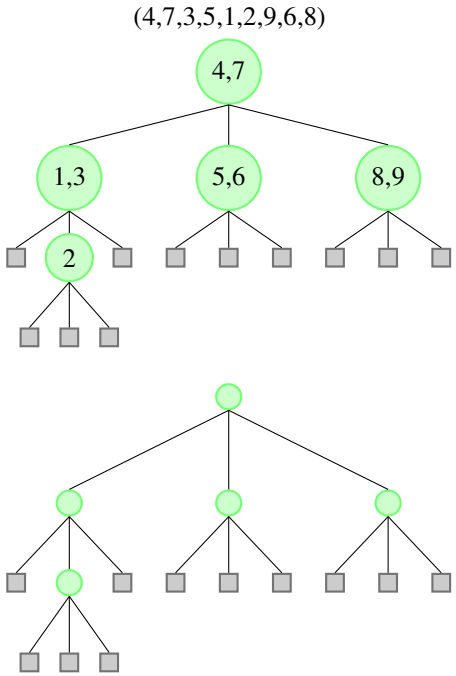

Fig. 1: Example of a 3-ary search tree of size 9 build from permutation $(4, 7, 3, 5, 1, 2, 9, 6, 8)$ and its unlabeled counterpart.

The standard probability model assumes that every permutation of the keys $\{1, \ldots, n\}$ is equally likely. The choice of pivots can then be deterministic. For example, one always chooses the first $m-1$ keys. However, after removing keys from a m-ary search tree we obtain an unlabeled search tree, but this time the probability distribution is non-uniform.

B. Unlabeled d-ary Plane Trees Generation

We consider the following generation model of an unlabeled d-ary plane tree. Suppose that the process starts with an empty tree, that is with just an external node (leaf). The first step in the growth process is to replace this external node by an internal one with d successors that are external nodes (see Figure 2). Then with probability $\frac{1}{d}$ one of these d external nodes is selected and again replaced by an internal node with d successors. In each next step one of the external nodes is replaced (with equal probability) by an internal node with d

successors. At the end, we remove the labels from internal nodes of a tree.*

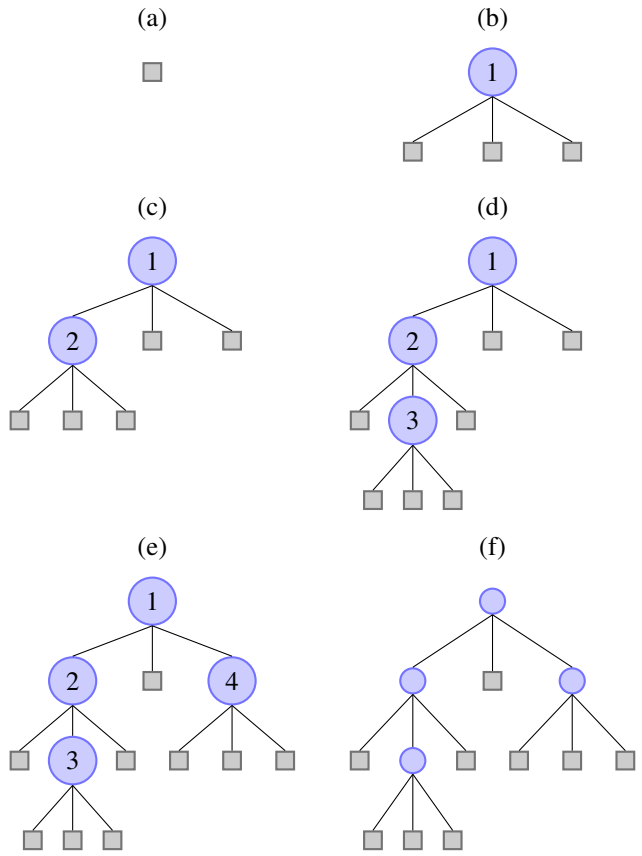

Fig. 2: Example of the generation process that produces 3-ary plane tree of size 4 and its unlabeled counterpart.

It is known that such evolution process (without removing the labels at the end) produces d-ary *recursive plane trees* (see [8]). By definition, d-ary plane recursive trees are rooted plane trees with labels on internal nodes and where each node has exactly d successors (either internal or external). The root is labeled by 1 and the labels of all internal successors of any node v are larger than the label of v. We define the size of the d-ary recursive plane tree by the number of internal nodes.

Let $\mathcal{F}$ denote a set of all unlabeled d-ary plane oriented trees and, for each integer $n \geq 0$, let $\mathcal{F}_n$ be a subset of $\mathcal{F}$ consisting of all trees that contains exactly n internal nodes. Let F_n be a random variable supported on $\mathcal{F}_n$. Similarly, let $\mathcal{G}$ be a set of all d-ary plane oriented recursive trees and, for each integer $n \geq 0$, let $\mathcal{G}_n$ be a subset of $\mathcal{G}$ consisting of all trees that contains exactly n internal nodes. Throughout the paper, we will denote a given unlabeled d-ary plane tree of size n as $\mathtt{f}_n$ and we will denote a given d-ary plane recursive tree of size n as $\mathtt{g}_n$.

*Observe that labels describe the history of the evolution process.

Let us formally define a source that generates unlabeled d-ary plane trees of size n as:

1) draw a d-ary plane recursive tree $\mathtt{g}_n$ uniformly from the set $\mathcal{G}_n$,
2) return a tree $\mathtt{f}_n \in \mathcal{F}_n$ by removing labels from $\mathtt{g}_n$.

The natural probability distribution on d-ary plane oriented recursive trees of size n is to assume that each of $|\mathcal{G}_n|$ trees is equally likely. Observe that the above described evolution process (choosing an external node to replace with uniform distribution among all external nodes) generates every d-ary plane recursive tree of size n in a unique way and with uniform distribution. However, after removing labels from a d-ary plane recursive tree $\mathtt{g}_n$ we obtain an unlabeled d-ary plane tree $\mathtt{f}_n$ that are generated non-uniformly. For example, there are $\binom{3}{2,0,1}\binom{1}{0,1,0} = 3$ ways to obtain resulting tree form Figure 2, but there is only one way (exactly $\binom{3}{3,0,0}\binom{2}{2,0,0}\binom{1}{1,0,0} = 1$) to obtain a tree where every internal node is a left most child of a parent node.

C. Unlabeled General Plane Trees Generation

We consider the following generation model of unlabeled plane trees. Suppose that the process starts with the root node carrying a label 1. Then we add a node with label 2 to the root. The next step is to attach a node with label 3. However, there are three possibilities: either to add it to the root (as a left or right successor) or to the node with label 2. Similarly one proceeds further. Now if a node already has out-degree k (where the descendants are ordered), then there are $k+1$ possible ways to add new node (this time we do not distinguish between external and internal nodes). Hence, if a plane tree already has $j-1$ nodes then there are precisely $2j-3$ possibilities to attach the j'th node (see Figure 3). More precisely, the probability of choosing a node of out-degree k equals $(k+1)/(2j-3)$. At the end, we remove the labels from internal nodes of a tree.

It is known that such evolution process (without removing the labels at the end) produces plane recursive trees (see [8]). By definition, plane recursive trees are rooted labeled plane trees, where the root is labeled by 1 and the labels of all successors of any node v are larger than label of v. We describe the size of the plane recursive tree by the number of nodes.

Let $\mathcal{T}$ denote a set of all unlabeled plane oriented trees and, for each integer $n \geq 0$, let $\mathcal{T}_n$ be a subset of $\mathcal{T}$ consisting of all trees that contains exactly n nodes and let $\mathcal{T}_n^{(d)}$ be a subset of $\mathcal{T}_n$ consisting of all trees that contains exactly n nodes and root degree equal to d. Let T_n be a random variable supported on $\mathcal{T}_n$. Similarly, let $\mathcal{R}$ be a set of all plane oriented recursive trees and, for each integer $n \geq 0$, let $\mathcal{R}_n$ be a subset of $\mathcal{R}$ consisting of all trees that contains exactly n nodes. Throughout the paper, we will denote a given unlabeled plane

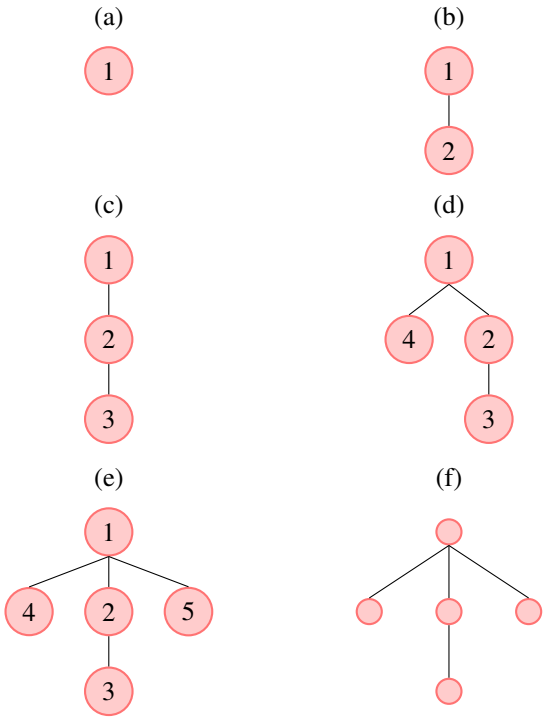

Fig. 3: Example of the generation process that produces unlabeled general plane tree of size 5 and its unlabeled counterpart.

tree of size n as $\mathtt{t}_n$ and we will denote a given plane recursive tree of size n as $\mathtt{r}_n$.

The natural probability distribution on plane oriented recursive trees of size n is to assume that each of $|\mathcal{R}_n|$ trees is equally likely. Observe that the above described evolution process generates every plane recursive tree of size n in a unique way and with uniform distribution. However, after removing labels from a plane recursive tree $\mathtt{r}_n$ we obtain an unlabeled plane tree $\mathtt{t}_n$, but this time the probability distribution on a set $\mathcal{T}_n$ is non-uniform. Observe that for instance there are $\binom{4}{1,2,1} = 12$ ways to obtain resulting tree form Figure 3, but there is only one way to obtain a tree where every node has exactly one successor.

III. Main Results

In this section we present our main results. In particular, we briefly address the entropy of m-ary search tree. Then we present our derivation of the recurrence for the entropy of d-ary trees that leads us to a formula for the entropy rate. Finally, we derive the entropy rate for general trees.

We should point out that in all our models, the probability of a tree generation is non-uniform and conditionally independent. Indeed, let T_n be a random variable representing a tree t_n on n internal nodes. Assume now that at the root we split t_n into d subtrees of size $k_1, \ldots, k_d$, respectively, where $k_1 + \cdots + k_d = n - 1$. Then the probability $P(T_n = t_n)$ of generating tree t_n in all our models satisfies

$$P(T_n = t_n) = P(k_1, \ldots, k_d) \prod_{i=1}^{d} P(T_{k_i} = t_{k_i}) \tag{3}$$

where $P(k_1, \ldots, k_d)$ is the probability of a split at the root of n internal nodes into subtrees $t_{k_1}, \ldots, t_{k_d}$, respectively. This split probability is different for m-ary search trees, d-ary trees, and general trees, as we shall see in this section.

Throughout we shall use the following notation. Let $\mathbf{k}^{(n)} = (k_1, \ldots, k_n)$ denote a n-dimensional vector and $\|\mathbf{k}^{(n)}\| = k_1 + \ldots + k_m$ be its L^1 norm. Let $(k, \mathbf{k}^{(n-1)}) = (k, k_2 \ldots, k_n)$ denote a n-dimensional vector with the first coordinate equal to k. We will write $\mathbf{k}$ instead of $\mathbf{k}^{(n)}$ when the vector dimension is obvious.

A. The Entropy of the Unlabeled m-ary Search Trees

Let U_n denote a random variable representing unlabeled m-ary search tree with n keys. We write $\mathtt{u}_n$ for a m-ary (unlabeled) search tree with n keys.

We describe now the splitting at the root for the search tree and denote this process by the random vector $\mathbf{Y}_n^{(m)} = (Y_{n,1}, \ldots, Y_{n,m})$, where $Y_{n,j} = |I_j|$ is the number of keys in the j-th subset. If $n \geq m - 1$ we have $Y_{n,1} + \cdots Y_{n,m} = n - m + 1$ and

$$\mathbb{P}\left(\mathbf{Y}_n^{(m)} = \mathbf{k}^{(m)}\right) = \frac{1}{\binom{n}{m-1}}. \tag{4}$$

Notice that $\mathbb{P}\left(\mathbf{Y}_n^{(m)} = \mathbf{k}^{(m)}\right)$ does not depends on the vector $\mathbf{k}^{(m)}$ coordinates $k_1, \ldots, k_m$, which simplifies calculations.

Suppose that the tree $\mathtt{u}_n$ has subtrees $\mathtt{u}_{k_1}, \ldots, \mathtt{u}_{k_m}$ of sizes $k_1, \ldots, k_m$, then by (3)

$$\mathbb{P}\left(U_n = \mathtt{u}_n\right) = \mathbb{P}\left(\mathbf{Y}_n^{(m)} = \mathbf{k}^{(m)}\right) \prod_{j=1}^{m} \mathbb{P}\left(U_{k_j} = \mathtt{u}_{k_j}\right). \tag{5}$$

Let us establish the initial conditions of the entropy of m-ary search trees. If $n = 0$ we have empty tree and the $H(U_0) = 0$. Moreover, if $1 \leq n \leq m-1$ all keys are stored in one node and $H(U_n) = 0$. For for $n > m - 1$ there is a bijection between a tree U_n and a tuple $(\mathbf{Y}_n^{(m)}, U_{Y_{n,1}}, \ldots, U_{Y_{n,m}})$ which is an immediate consequence of (5). Therefore, for $n > m - 1$, we have

$$\begin{aligned} H(U_n) &= H\left(\mathbf{Y}_n^{(m)}, U_{Y_{n,1}}, \ldots, U_{Y_{n,m}}\right) \\ &= H\left(\mathbf{Y}_n^{(m)}\right) + H\left(U_{Y_{n,1}}, \ldots, U_{Y_{n,m}} | \mathbf{Y}_n^{(m)}\right) \\ &= H\left(\mathbf{Y}_n^{(m)}\right) + \\ &+ \sum_{\|\mathbf{k}\| = n-m+1} H\left(U_{Y_{n,1}}, \ldots, U_{Y_{n,m}} | \mathbf{Y}_n^{(m)} = \mathbf{k}^{(m)}\right) \\ &\cdot \mathbb{P}\left(\mathbf{Y}_n^{(m)} = \mathbf{k}^{(m)}\right). \end{aligned}$$

Observe that $H\left(U_{Y_{n,1}},\dots,U_{Y_{n,m}}|\mathbf{Y}_n^{(m)}=\mathbf{k}^{(m)}\right) = H\left(U_{k_1},\dots,U_{k_m}\right)$. By independence of subtrees $U_{k_1},\dots,U_{k_m}$ we have

$$H\left(U_n\right) = H\left(\mathbf{Y}_n^{(m)}\right) + \\ + m\sum_{k=1}^{n-m+1} H\left(U_k\right) \sum_{\|\mathbf{k}^{(m-1)}\|=n-m+1-k} \mathbb{P}\left(\mathbf{Y}_n^{(m)}=(k,\mathbf{k}^{(m-1)})\right) \\ = H\left(\mathbf{Y}_n^{(m)}\right) + m\sum_{k=0}^{n-m+1} H\left(U_k\right)\mathbb{P}\left(Y_{n,1}=k\right).$$

For $n\geq m-1$ and $1\leq j\leq m$, the components $Y_{n,j}$ are identically distributed, and for $0\leq k\leq n-1$,

$$\mathbb{P}\left(Y_{n,1}=k\right) = \sum_{\|\mathbf{k}^{(m-1)}\|=n-m+1-k} \mathbb{P}\left(\mathbf{Y}_n^{(m)}=\left(k,\mathbf{k}^{(m-1)}\right)\right) \\ = \frac{\binom{n-k-1}{m-2}}{\binom{n}{m-1}}. \tag{6}$$

The last equation comes from the fact that there are $\binom{n-l-1}{m-2}$ ways of split $n-l-1$ keys into $m-1$ sublists (i.e. choosing $m-2$ pivots from $n-l-1$ keys) (it also can be found e.g. in [8]).

On the other hand, from (4) we have

$$H\left(\mathbf{Y}_n^{(m)}\right) = \\ -\sum_{\|\mathbf{k}\|=n-m+1} \mathbb{P}\left(\mathbf{Y}_n^{(m)}=\mathbf{k}^{(m)}\right)\log\mathbb{P}\left(\mathbf{Y}_n^{(m)}=\mathbf{k}^{(m)}\right) \\ = \frac{1}{\binom{n}{m-1}}\log\binom{n}{m-1}\sum_{\|\mathbf{k}\|=n-m+1} 1 = \log\binom{n}{m-1}.$$

The last equality comes from the fact that the sum $\sum_{\|\mathbf{k}\|=n-m+1}1$ equals to the number of choices of $m-1$ pivots from n keys, which is $\binom{n}{m-1}$.

Finally, we arrive at the following recurrence for the entropy

$$H\left(U_n\right) = \log\binom{n}{m-1} + \frac{m}{\binom{n}{m-1}}\sum_{k=0}^{n-m+1}\binom{n-k-1}{m-2}H\left(U_k\right).$$

The asymptotic of a recurrence like above was studied before; see Proposition 7 in [5] and Theorem 2.4 in [9] that we quote below.

Theorem 1 ([9], Theorem 2.4, Asymptotic Transfer Theorem)**.** *Let*

$$a_n = b_n + \frac{m}{\binom{n}{m-1}}\sum_{j=0}^{n-(m-1)}\binom{n-1-j}{m-2}a_j, \quad n\geq m-1,$$

with specified initial conditions (say) $a_j := b_j$, $0\leq j\leq m-2$. If

$$b_n = o\left(n\right) \quad \textit{and} \quad \sum_{n\geq 0}\frac{b_n}{(n+1)(n+2)} \textit{ converges,}$$

then

$$a_n = \frac{K_1}{\mathcal{H}_m-1}n + o\left(n\right), \quad \textit{where} \quad K_1 := \sum_{j\geq 0}\frac{b_j}{(j+1)(j+2)}.$$

Hence, the entropy of the m-ary search tree becomes

$$H(U_n) = c_m n + o\left(n\right),$$

where

$$c_m = 2\phi_m\sum_{k\geq 0}\frac{\log\binom{k}{m-1}}{(k+1)(k+2)}$$

and $\phi_m = \frac{1}{2\mathcal{H}_m-2}$ is called occupancy constant.

Observe that if $m=2$ the number of nodes of m-ary search tree equals n, but for $m>2$ the number of nodes of m-ary search tree is a random variable $S_{n,m}$. Knuth [12] was the first to show that $\mathbb{E}\left(S_{n,m}\right)\sim\phi_m n$.

In order to compare constant c_m of m-ary search tree with d-ary trees discussed next we should point out that m-ary search trees of size n have on average $\sim\phi_m n$ internal nodes. Thus, it makes sense to normalize the constant c_m as $\hat{c}_m = c_m/\phi_m$. Then numerically $\hat{c}_2\approx 1.73638, \hat{c}_3\approx 2.5014, \hat{c}_4\approx 2.93994, \hat{c}_5\approx 3.22688$.

Using Theorem 1 and the fact that $H\left(\mathbf{Y}_n^{(m)}\right) = \log\binom{n}{m-1} = o\left(n\right)$, we conclude this section with the following result.

Corollary 1. *The entropy rate $h_m(u) = \lim_{n\to\infty}H\left(U_n\right)/n$ of the unlabeled m-ary trees, generated according to the m-ary search trees model, is given by*

$$h_m(u) = 2\phi_m\sum_{k\geq 0}\frac{\log\binom{k}{m-1}}{(k+1)(k+2)}. \tag{7}$$

Remark 1. It can be checked numerically that the entropy rate of the unlabeled m-ary search trees, generated according to the model of m-ary search trees for $m=2,3,4,5$ is $h_2(u)\approx 1.73638, h_3(u)\approx 1.50084, h_4(u)\approx 1.3569, h_5(u)\approx 1.25723$ respectively.

B. The Entropy of the Unlabeled d-ary Plane Trees

Let $g_n = |\mathcal{G}_n|$ be the number of d-ary plane recursive trees with n internal nodes. From [8] we know that for $d=2$ we have $g_n = n!$. Moreover, for $d>2$ we have

$$g_n = (-1)^n(d-1)^n\frac{\Gamma(2-\frac{d}{d-1})}{\Gamma(2-\frac{d}{d-1}-n)}. \tag{8}$$

Let $\mathcal{G}_{\mathtt{f}_n}$ denote a subset of $\mathcal{G}_n$ of trees that have the same structure as a tree $\mathtt{f}_n\in\mathcal{F}_n$, that is, an unlabeled version of a tree $\mathtt{g}_n\in\mathcal{G}_n$ is isomorphic to $\mathtt{f}_n$. Moreover, let $g_{\mathtt{f}_n} = |\mathcal{G}_{\mathtt{f}_n}|$ be the number of d-ary plane recursive trees that have the same structure as a tree $\mathtt{f}_n$. Observe that the probability that above defined source generates a given unlabeled tree $\mathtt{f}_n\in\mathcal{F}_n$ is

$$\mathbb{P}\left(F_n=\mathtt{f}_n\right) = \frac{g_{\mathtt{f}_n}}{g_n}. \tag{9}$$

Suppose that the tree $\mathtt{f}_n$ has subtrees $\mathtt{f}_{k_1}, \ldots, \mathtt{f}_{k_d}$ of sizes $k_1, \ldots, k_d$. Then

$$\begin{aligned}\mathbb{P}\left(F_n = \mathtt{f}_n\right) &= \frac{1}{g_n}\binom{n-1}{k_1, \ldots, k_d}\prod_{j=1}^{d} g_{\mathtt{f}_{k_j}} \\ &= \binom{n-1}{k_1, \ldots, k_d}\frac{g_{k_1}\cdots g_{k_d}}{g_n}\prod_{j=1}^{d}\mathbb{P}\left(F_{k_j} = \mathtt{f}_{k_j}\right).\end{aligned} \tag{10}$$

Observe that $\binom{n-1}{k_1,\ldots,k_d}\frac{g_{k_1}\cdots g_{k_d}}{g_n}$ is the probability that the subtrees of the root are of sizes $k_1, \ldots, k_d$. Let us define a random vector $\mathbf{V}_n^{(d)} : \mathcal{G}_n \to \{0, \ldots, n-1\}^d$, where its j'th component $V_{n,j}$ denotes the size of j-th subtree. For $n \geq 1$ we have $V_{n,1} + \ldots + V_{n,d} = n-1$ and

$$\mathbb{P}\left(\mathbf{V}_n^{(d)} = \mathbf{k}^{(d)}\right) = \binom{n-1}{k_1, \ldots, k_d}\frac{g_{k_1}\cdots g_{k_d}}{g_n}. \tag{11}$$

Define the entropy of unlabeled d-ary plane tree of size n as

$$H\left(F_n\right) = -\sum_{\mathtt{f}_n \in \mathcal{F}_n}\mathbb{P}\left(F_n = \mathtt{f}_n\right)\log\left(\mathbb{P}\left(F_n = \mathtt{f}_n\right)\right).$$

Let us establish the initial conditions of the entropy of our source. If $n = 0$ we have empty tree, and $H\left(F_0\right) = 0$. If $n = 1$, we have one fixed tree and $H\left(F_1\right) = 0$. By (10) for $n > 1$ there is a bijection between a tree F_n and a tuple $(\mathbf{V}_n^{(d)}, F_{V_1}, \ldots, F_{V_d})$. Therefore, for $n > 1$, we have

$$\begin{aligned}H\left(F_n\right) &= H\left(\mathbf{V}_n^{(d)}, F_{V_{n,1}}, \ldots, F_{V_{n,d}}\right) \\ &= H\left(\mathbf{V}_n^{(d)}\right) + H\left(F_{V_{n,1}}, \ldots, F_{V_{n,d}} | \mathbf{V}_n^{(d)}\right) \\ &= H\left(\mathbf{V}_n^{(d)}\right) + \\ &\quad \sum_{\|\mathbf{k}\| = n-1} H\left(F_{k_1}, \ldots, F_{k_d}\right)\mathbb{P}\left(\mathbf{V}_n^{(d)} = \mathbf{k}^{(d)}\right).\end{aligned}$$

Since subtrees $F_{k_1}, \ldots, F_{k_d}$ are (conditionally) independent, we have

$$\begin{aligned}H\left(F_n\right) = H\left(\mathbf{V}_n^{(d)}\right) + \\ d\sum_{k=0}^{n-1} H\left(F_k\right) \sum_{\|\mathbf{k}^{(d-1)}\| = n-1-k} \mathbb{P}\left(\mathbf{V}_n^{(d)} = \left(k, \mathbf{k}^{(d-1)}\right)\right).\end{aligned}$$

For $k = 0, \ldots, n-1$, let $p_{n,k}$ be the probability that one specified subtree in a d-ary recursive tree is of size k, that is,

$$p_{n,k} = \sum_{\|\mathbf{k}^{(d-1)}\| = n-1-k} \mathbb{P}\left(\mathbf{V}_n^{(d)} = \left(k, \mathbf{k}^{(d-1)}\right)\right). \tag{12}$$

Then

$$H\left(F_n\right) = H\left(\mathbf{V}_n^{(d)}\right) + d\sum_{k=0}^{n-1} H\left(F_k\right) p_{n,k}. \tag{13}$$

Lemma 1. *For $k = 0, \ldots, n-1$ and $d > 1$, let $\alpha = \frac{d}{d-1}$, then*

$$p_{n,k} = \frac{(\alpha-1)}{n}\frac{n!\Gamma(k+\alpha-1)}{k!\Gamma(n+\alpha-1)}.$$

Proof. Using (11), we can rewrite (12) as

$$\begin{aligned}p_{n,k} = \frac{(n-1)!g_k}{k!(n-1-k)!g_n} \times \\ \sum_{k_2+\ldots+k_d = n-1-k}\binom{n-1-k}{k_2, \ldots, k_d} g_{k_2}\cdots g_{k_d}.\end{aligned}$$

Let us define the exponential generating function $G(z) = \sum_{n\geq 0} g_n \frac{z^n}{n!}$ with $g_0 = 1$. From [8] we know that

$$G(z) = (1-(d-1)z)^{-\frac{1}{d-1}}.$$

Observe that

$$\sum_{k_2+\ldots+k_d = n-1-k}\binom{n-1-k}{k_2, \ldots, k_d} g_{k_2}\cdots g_{k_d}$$

is the $n-1-k$'th coefficient of the function $G(z)^{d-1}$ (denoted as $\left[\frac{z^{n-1-k}}{(n-1-k)!}\right]G(z)^{d-1}$). Hence

$$\begin{aligned}p_{n,k} &= \frac{(n-1)!g_k}{k!(n-1-k)!g_n}\left[\frac{z^{n-1-k}}{(n-1-k)!}\right]G(z)^{d-1} \\ &= \frac{(n-1)!g_k}{k!g_n}\left[z^{n-1-k}\right]\frac{1}{1-(d-1)z} \\ &= \frac{(n-1)!g_k}{k!g_n}(d-1)^{n-1-k}.\end{aligned}$$

For $d = 2$, we have $g_n = n!$ and the result is immediate. For $d > 2$, from (8) we find

$$p_{n,k} = \frac{(\alpha-1)}{n}\frac{(-1)^n n!\Gamma(2-\alpha-n)}{(-1)^k k!\Gamma(2-\alpha-k)}.$$

From [19] we know that $\Gamma(z-n) = \frac{(-1)^n\pi}{\Gamma(n+1-z)\sin(\pi z)}$, hence

$$(-1)^n\Gamma(n+\alpha)\Gamma\left(2-\alpha-n\right) = \frac{\pi(n-1+\alpha)}{\sin(\pi(2-\alpha))}, \tag{14}$$

and then

$$p_{n,k} = \frac{(\alpha-1)}{n}\frac{n!\Gamma(k+\alpha)(n+\alpha-1)}{k!\Gamma(n+\alpha)(k+\alpha-1)}.$$

Since $\Gamma(z+1) = z\Gamma(z)$ we get desired result. □

Remark 2. Observe that for $d = 2$, we have $\alpha = 2$ and $p_{n,k} = \frac{1}{n}$. It does not depend on k, which greatly simplifies computations as shown in [13]. Moreover, it equals $\mathbb{P}\left(Y_{n,1} = k\right)$ in the case of the binary search trees. Therefore, two models for 2-ary plane trees and for binary search trees are equal. For $d > 2$ it is not the case. For instance for $d = 3$, we have $\alpha = \frac{3}{2}$ and

$$p_{n,k} = \frac{1}{2n}\frac{\binom{2k}{k}2^{2n}}{\binom{2n}{n}2^{2k}},$$

which clearly depends on k and does not equal $\frac{n-k-1}{\binom{n}{2}}$ what would be the in case of 3-ary search trees.

The recurrence presented in (13) is a novel one that we need to solve. In lemma below we propose a general solution of recurrences of this form.

Lemma 2. *For constant α, x_0 and x_1, the recurrence*

$$x_n = a_n + \frac{\alpha}{n}\frac{n!}{\Gamma(n+\alpha-1)}\sum_{k=0}^{n-1}\frac{\Gamma(k+\alpha-1)}{k!}x_k, \qquad n \geq 2 \tag{15}$$

has the following solution for $n \geq 2$

$$x_n = a_n + \alpha(n+\alpha-1)\sum_{k=0}^{n-1}\frac{a_k}{(k+\alpha-1)(k+\alpha)} + \frac{n+\alpha-1}{\alpha+1}\left(x_1 + \frac{x_0}{\alpha-1}\right).$$

Proof. Let us divide both sides of the recurrence by the normalizing factor $\frac{\Gamma(n+\alpha-1)}{n!}$. Define also $\hat{x}_n = \frac{x_n\Gamma(n+\alpha-1)}{n!}$ and $\hat{a}_n = \frac{a_n\Gamma(n+\alpha-1)}{n!}$. Then

$$\hat{x}_n = \hat{a}_n + \frac{\alpha}{n}\sum_{k=2}^{n-1}\hat{x}_k. \tag{16}$$

To solve recurrence (16) we compute $n\hat{x}_n - (n-1)\hat{x}_{n-1}$. This leads us to

$$\hat{x}_n = \hat{a}_n - \left(1-\frac{1}{n}\right)\hat{a}_{n-1} + \left(1+\frac{\alpha-1}{n}\right)\hat{x}_{n-1},$$

that holds for $n \geq 3$. Then after iterating the above we arrive at

$$\hat{x}_n = \hat{x}_2\prod_{j=3}^{n}\left(1+\frac{\alpha-1}{j}\right) + \sum_{k=3}^{n}\left(\hat{a}_k - \left(1-\frac{1}{k}\right)\hat{a}_{k-1}\right)\prod_{j=k+1}^{n}\left(1+\frac{\alpha-1}{j}\right). \tag{17}$$

The product $\prod_{j=k+1}^{n}\left(1+\frac{\alpha-1}{j}\right) = \frac{k!\Gamma(n+\alpha)}{n!\Gamma(k+\alpha)}$, and after some standard calculations we obtain

$$\hat{x}_n = \hat{a}_n + (\hat{x}_2 - \hat{a}_2)\frac{2\Gamma(n+\alpha)}{\Gamma(\alpha+2)n!} + \frac{\Gamma(n+\alpha)}{n!}\sum_{k=2}^{n-1}\hat{a}_k\frac{k!}{\Gamma(k+\alpha)}\frac{\alpha}{k+\alpha}.$$

Going back from $\hat{x}_n$ and $\hat{a}_n$ to x_n, a_n, respectively, we obtain

$$x_n = a_n + \alpha(n+\alpha-1)\sum_{k=2}^{n-1}\frac{a_k}{(k+\alpha-1)(k+\alpha)} + (x_2 - a_2)\frac{n+\alpha-1}{\alpha+1}.$$

Observe that $x_2 - a_2 = x_1 + \frac{x_0}{\alpha-1}$. This completes the proof. □

Remark 3. Observe that for $d = 2$ and $x_0 = x_1 = 0$ and $a_n = o(n)$, we have $\alpha = 2$ and

$$x_n = 2n\sum_{k\geq 0}\frac{a_k}{(k+1)(k+2)} + o(n)$$

as in [13]. But with the same assumptions and $d = 3$ we have

$$x_n = \frac{3}{2}n\sum_{k\geq 0}\frac{a_k}{(k+\frac{1}{2})(k+\frac{3}{2})} + o(n).$$

This leads us to out first main result.

Theorem 2. *The entropy of an unlabeled d-ary plane tree, generated according to the model of d-ary plane recursive tree, is given by*

$$H(F_n) = H\left(\mathbf{V}_n^{(d)}\right) + \alpha(n+\alpha-1)\sum_{k=0}^{n-1}\frac{H\left(\mathbf{V}_k^{(d)}\right)}{(k+\alpha-1)(k+\alpha)}, \tag{18}$$

where $\alpha = \frac{d}{d-1}$ and

$$H\left(\mathbf{V}_n^{(d)}\right) = -\sum_{\|\mathbf{k}\|=n-1}\mathbb{P}\left(\mathbf{V}_n^{(d)} = \mathbf{k}^{(d)}\right)\log\mathbb{P}\left(\mathbf{V}_n^{(d)} = \mathbf{k}^{(d)}\right).$$

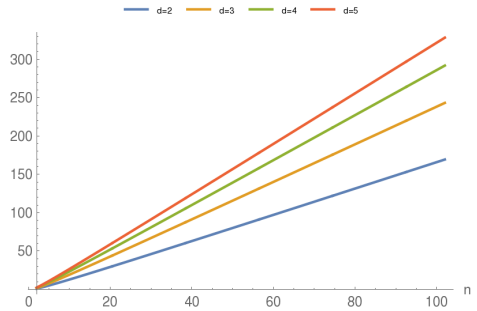

Fig. 4: Entropy of d-ary plane trees for $d = 2, 3, 4, 5$ respectively and number of nodes $n \in [1, 100]$.

We conclude this section with the following result.

Lemma 3. *The entropy rate $h_d(f) = \lim_{n\to\infty} H(F_n)/n$ of the unlabeled d-ary plane trees, generated according to the model of d-ary plane recursive trees, is given by*

$$h_d(f) = \alpha\sum_{k=0}^{\infty}\frac{H(\mathbf{V}_k)}{(k+\alpha-1)(k+\alpha)}, \tag{19}$$

with $\alpha = \frac{d}{d-1}$.

Proof. Having Theorem 2 in mind, we just need to prove that $H\left(\mathbf{V}_n^{(d)}\right) = o(n)$. Let us recall that $\mathbf{V}_n^{(d)} : \mathcal{G}_n \to \{0, \ldots, n-1\}^d$. Since the entropy of random variable is upper bounded by the logarithm of the variable image cardinality, we have

$$H\left(\mathbf{V}_n^{(d)}\right) \leq \log\left(n^d\right) = o(n).$$

□

Remark 4. Taking a closer look at $H\left(\mathbf{V}_n^{(d)}\right)$ we find

$$H\left(\mathbf{V}_n^{(d)}\right) = \log\left(n\frac{g_n}{n!}\right) - d\sum_{k=0}^{n-1} p_{n,k}\log\left(\frac{g_k}{k!}\right).$$

In particular, $H\left(\mathbf{V}_n^{(2)}\right) = H\left(\mathbf{Y}_n^{(2)}\right) = \log(n)$ and the entropy rate $h_2(f) \approx 1.73638$, which matches the entropy rate of the binary search trees. On the other hand, for $d = 3$ and $n > 0$ we have

$$H\left(\mathbf{V}_n^{(3)}\right) = \log\left(\frac{n}{2^n}\binom{2n}{n}\right) - \frac{3}{2n}\sum_{k=0}^{n-1}\frac{\binom{2k}{k}2^{2n}}{\binom{2n}{n}2^{2k}}\log\left(\frac{\binom{2k}{k}}{2^k}\right).$$

This allows us to check numerically that the entropy rate of the unlabeled 3-ary plane trees, generated according to the model of 3-ary recursive trees is $h_3(u) \approx 2.470$.

C. The Entropy of the Unlabeled General Plane Trees

Let $r_n = |\mathcal{R}_n|$. From [8] we know that there are

$$r_n = (2n-3)!! = \frac{n!}{n2^{n-1}}\binom{2n-2}{n-1} \tag{20}$$

different plane oriented recursive trees of size n.

As in the case of the d-ary plane trees, let $\mathcal{R}_{\mathtt{t}_n}$ denote a subset of $\mathcal{R}_n$ trees that have the same structure as a tree $\mathtt{t}_n \in \mathcal{T}_n$; moreover, let $r_{\mathtt{t}_n} = |\mathcal{R}_{\mathtt{t}_n}|$ be the number of plane recursive trees that have the same structure as a tree $\mathtt{t}_n$. Observe that

$$\mathbb{P}\left(T_n = \mathtt{t}_n\right) = \frac{r_{\mathtt{t}_n}}{r_n}. \tag{21}$$

Let D_n denote a random variable representing the number of subtrees of the root. Observe that $\mathbb{P}\left(D_n = d\right) = \frac{r_n^{(d)}}{r_n}$, where $r_n^{(d)} = |\mathcal{R}_n^{(d)}|$ is the number of plane recursive trees with the root degree equal d. Suppose that the tree $\mathtt{t}_n$ has d subtrees $\mathtt{t}_{k_1}, \ldots, \mathtt{t}_{k_d}$ of sizes $k_1, \ldots, k_d$. Then

$$\begin{aligned}
&\mathbb{P}\left(T_n = \mathtt{t}_n \;\&\; \mathtt{t}_n \text{ root degree equals } d\right)\\
&\quad= \mathbb{P}\left(D_n = d\right)\mathbb{P}\left(T_n = \mathtt{t}_n | D_n = d\right)\\
&\quad= \binom{n-1}{k_1,\ldots,k_d}\frac{r_{k_1}\cdots r_{k_d}}{r_n}\prod_{j=1}^{d}\mathbb{P}\left(T_{k_j} = \mathtt{t}_{k_j}\right).
\end{aligned} \tag{22}$$

Observe that $\binom{n-1}{k_1,\ldots,k_d}\frac{r_{k_1}\cdots r_{k_d}}{r_n}$ is a probability that the root of plane recursive tree of size n has degree equal to d and the root's subtrees are of sizes $k_1, \ldots, k_d$. Let $\mathbf{W}_n^{(d)} : \mathcal{R}_n^{(d)} \to \{1, \ldots, n-d\}^d$, where its j'th component $W_{n,j}$ denotes the size of j-th subtree when the root is of degree d. For $n \geq 1$ we have $W_{n,1} + \ldots + W_{n,d} = n-1$ and

$$\begin{aligned}
\mathbb{P}\left(D_n = d\right)\mathbb{P}\left(\mathbf{W}_n^{(D_n)} = \mathbf{k}^{(D_n)} | D_n = d\right) =\\
\binom{n-1}{k_1,\ldots,k_d}\frac{r_{k_1}\cdots r_{k_d}}{r_n}.
\end{aligned} \tag{23}$$

Let us define the entropy of unlabeled plane tree of size n as

$$H\left(T_n\right) = -\sum_{\mathtt{t}_n \in \mathcal{T}_n}\mathbb{P}\left(T_n = \mathtt{t}_n\right)\log\left(\mathbb{P}\left(T_n = \mathtt{t}_n\right)\right).$$

Observe that

$$\begin{aligned}
H\left(T_n, D_n\right) = H\left(T_n\right) -\\
\sum_{\substack{\mathtt{t}_n \in \mathcal{T}_n\\ 0 < d < n}}\mathbb{P}\left(T_n = \mathtt{t}_n, D_n = d\right)\log\mathbb{P}\left(D_n = d | T_n = \mathtt{t}_n\right),
\end{aligned}$$

Since D_n is a function of the tree random variable T_n we have $\mathbb{P}\left(D_n = d | T_n = \mathtt{t}_n\right) = 1$. Therefore, $H\left(T_n, D_n\right) = H\left(T_n\right)$ and in order to calculate $H\left(T_n\right)$ we can use (22).

Let us establish the initial conditions of the entropy of our source. If $n = 1$ we have just a root node, and the $H\left(T_1\right) = 0$, similarly if $n = 2$, we have one fixed tree and the $H\left(T_2\right) = 0$. Let us observe that for $n > 2$ and tree root's degree equals to d, there is a bijection between a tree T_n and a tuple $(\mathbf{W}_n^{(d)}, T_{W_{n,1}}, \ldots, T_{W_{n,d}})$ which is an immediate consequence of (22). Therefore, for $n > 2$, we have

$$\begin{aligned}
H\left(T_n\right) &= \sum_{d=1}^{n-1} H\left(\mathbf{W}_n^{(d)}, T_{W_{n,1}}, \ldots, T_{W_{n,d}}\right)\\
&= \sum_{d=1}^{n-1}\left(H\left(\mathbf{W}_n^{(d)}\right) + H\left(T_{W_{n,1}}, \ldots, T_{W_{n,d}} | \mathbf{W}_n^{(d)}\right)\right)\\
&= \sum_{d=1}^{n-1} H\left(\mathbf{W}_n^{(d)}\right) +\\
&\quad \sum_{d=1}^{n-1}\sum_{\|\mathbf{k}\|=n-1} H\left(T_{k_1}, \ldots, T_{k_d}\right)\mathbb{P}\left(\mathbf{W}_n^{(d)} = \mathbf{k}^{(d)}\right).
\end{aligned}$$

From conditional independence $T_{k_1}, \ldots, T_{k_d}$ we conclude

$$\begin{aligned}
H\left(T_n\right) &= \sum_{d=1}^{n-1} H\left(\mathbf{W}_n^{(d)}\right) +\\
&\sum_{d=1}^{n-1} d\sum_{k=1}^{n-d} H\left(T_k\right)\sum_{\|\mathbf{k}^{(d-1)}\|=n-1-k}\mathbb{P}\left(\mathbf{W}_n^{(d)} = \left(k, \mathbf{k}^{(d-1)}\right)\right).
\end{aligned}$$

For $k = 1, \ldots, n-1$, let $q_{n,k}$ be defined as the probability that one specified subtree in a plane recursive tree, with root's degree equals to d, is of size k. Hence

$$q_{n,k}^{(d)} = \sum_{\|\mathbf{k}^{(d-1)}\|=n-1-k}\mathbb{P}\left(\mathbf{W}_n^{(d)} = \left(k, \mathbf{k}^{(d-1)}\right)\right). \tag{24}$$

Therefore

$$H\left(T_n\right) = \sum_{d=1}^{n-1} H\left(\mathbf{W}_n^{(d)}\right) + \sum_{d=1}^{n-1} d\sum_{k=1}^{n-d} H\left(T_k\right) q_{n,k}^{(d)}. \tag{25}$$

We need an expression for the probability $q_{n,k}^{(d)}$ which we present in the next lemma.

Lemma 4. *For $k = 1, \ldots, n-1$ we have*

- $q^{(1)}_{n,n-1} = \frac{1}{2n-3}$ *and if* $k \neq n-1 : q^{(1)}_{n,k} = 0$,
- *for* $d > 1$*:*

$$q^{(d)}_{n,k} = 2^d \frac{d-1}{k(n-1-k)} \frac{\binom{2k-2}{k-1}\binom{2(n-1-k)-d}{n-2-k}}{\binom{2n-2}{n-1}}.$$

Proof. If $d = 1$ then the root has only 1 subtree with all other nodes, so its size has to be equal to $n-1$ and

$$q^{(1)}_{n,n-1} = \frac{r_{n-1}}{r_n} = \frac{1}{2n-3};$$

moreover, if $k \neq n-1 : q^{(1)}_{n,k} = 0$. In the case of $d > 1$, using (23), we can rewrite (24) as follows

$$q^{(d)}_{n,k} = \frac{(n-1)!r_k}{k!(n-1-k)!r_n} \times \sum_{k_2+\ldots+k_d=n-1-k} \binom{n-1-k}{k_2, \ldots, k_d} r_{k_2} \cdots r_{k_d}.$$

Let us define the exponential generating function $R(z) = \sum_{n\geq 0} r_n \frac{z^n}{n!}$ with $g_0 = 0$. Observe that

$$\sum_{k_2+\ldots+k_d=n-1-k} \binom{n-1-k}{k_2, \ldots, k_d} r_{k_2} \cdots r_{k_d}$$

is the $n-1-k$'th coefficient of the function $R(z)^{d-1}$ (denoted as $\left[\frac{z^{n-1-k}}{(n-1-k)!}\right] R(z)^{d-1}$). Therefore,

$$q_{n,k} = \frac{(n-1)!r_k}{k!r_n} \left[z^{n-1-k}\right] R(z)^{d-1}.$$

From (20) we find $R(z) = 1 - \sqrt{1-2z}$, which is also the solution of the equation

$$R = \frac{z}{1 - \frac{R}{2}}.$$

Hence, by Lagrange's inversion formula (see [10]), we obtain explicit formula for

$$[z^{n-1-k}]R(z)^{d-1} = 2^{d-n+k} \frac{d-1}{n-1-k} \binom{2(n-1-k)-d}{n-2-k}.$$

Putting everything together we arrive at the desired result. □

Recurrence found in (25) is another recurrence that we need to analyze. Its general solution is presented next.

Lemma 5. *For constant y_1 and y_2, the recurrence*

$$y_n = b_n + \sum_{d=1}^{n-1} d \sum_{k=1}^{n-d} q^{(d)}_{n,k} \cdot y_k, \qquad n > 2 \tag{26}$$

has the following solution for $n > 2$

$$y_n = \frac{2(2n-1)}{3} b_1 + b_n + \frac{1}{2}\left(n - \frac{1}{2}\right) \sum_{j=2}^{n-1} \frac{b_j}{\left(j - \frac{1}{2}\right)\left(j + \frac{1}{2}\right)}.$$

Proof. Using Lemma 4, for $n > 2$, we find

$$y_n = b_n + \frac{y_{n-1}}{2n-3} + \sum_{d=2}^{n-1} d(d-1)2^d \sum_{k=1}^{n-d} \frac{y_k}{k(n-1-k)} \frac{\binom{2k-2}{k-1}\binom{2n-2k-2-d}{n-k-2}}{\binom{2n-2}{n-1}}.$$

Multiplying both sides by $\frac{\binom{2n-2}{n-1}}{n}$ and substituting $\hat{y}_n = \frac{y_n\binom{2n-2}{n-1}}{n}$, $\hat{b}_n = \frac{b_n\binom{2n-2}{n-1}}{n}$ we get

$$\hat{y}_n = \hat{b}_n + \frac{2\hat{y}_{n-1}}{n(n-1)} + \frac{1}{n}\sum_{d=2}^{n-1} d(d-1)2^d \sum_{k=1}^{n-d} \frac{\hat{y}_k}{(n-1-k)} \binom{2n-2k-2-d}{n-k-2}.$$

Changing the order of summation gives us

$$\sum_{d=2}^{n-1} d(d-1)2^d \sum_{k=1}^{n-d} \frac{\hat{y}_k}{(n-1-k)} \binom{2n-2k-2-d}{n-k-2} = \sum_{j=1}^{n-2} \frac{\hat{y}_j}{n-j-1} \sum_{s=0}^{n-j} s(s-1)2^s \binom{2n-2j-2-s}{n-j-2}.$$

Since for $N > 0$:

$$\sum_{s=0}^{N} s(s-1)2^s \binom{2N-2-s}{N-2} = (N-1)2^{2N-1},$$

we obtain

$$\hat{y}_n = \hat{b}_n + \frac{2\hat{y}_{n-1}}{n(n-1)} + \frac{1}{n}\sum_{j=1}^{n-2} \hat{y}_j 2^{2n-2j-1}.$$

Dividing both sides by 2^{2n} and substituting $\tilde{y}_n = \frac{\hat{y}_n}{2^{2n}}$, $\tilde{b}_n = \frac{\hat{b}_n}{2^{2n}}$ we find

$$\tilde{y}_n = \tilde{b}_n + \frac{1}{2n}\sum_{j=1}^{n-1} \tilde{y}_j.$$

Solving this recurrence relation by calculating $n\tilde{y}_n - (n-1)\tilde{y}_{n-1}$ we obtain

$$\tilde{y}_n = b_1 \frac{\Gamma\left(n+\frac{1}{2}\right)}{\Gamma\left(\frac{5}{2}\right)n!} + \tilde{b}_n + \frac{\Gamma\left(n+\frac{1}{2}\right)}{n!} \sum_{j=2}^{n-1} \frac{\tilde{b}_j}{2j+1} \frac{j!}{\Gamma\left(j+\frac{1}{2}\right)}.$$

Substituting $\tilde{y}_n$ into y_n with $\tilde{y}_n = y_n \frac{\binom{2n-2}{n-1}}{n2^{2n}}$, we find the desired result. □

This leads us to our second main result.

Theorem 3. *The entropy of an unlabeled general plane tree, generated according to the model of plane recursive tree, is given by*

$$H(T_n) = \sum_{d=1}^{n-1} H\left(\mathbf{W}^{(d)}_n\right) + \frac{1}{2}\left(n - \frac{1}{2}\right) \sum_{j=2}^{n-1} \frac{\sum_{d=1}^{j-1} H\left(\mathbf{W}^{(d)}_k\right)}{\left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right)}, \tag{27}$$

where

$$H\left(\mathbf{W}_n^{(d)}\right) = -\sum_{\|\mathbf{k}\|=n-1} \mathbb{P}\left(\mathbf{W}_n^{(d)} = \mathbf{k}^{(d)}\right) \log \mathbb{P}\left(\mathbf{W}_n^{(d)} = \mathbf{k}^{(d)}\right).$$

We conclude this section with the following result.

Lemma 6. *The entropy rate $h(t) = \lim_{n\to\infty} H(T_n)/n$ of the unlabeled general plane trees, generated according to the model of plane recursive trees, is given by*

$$h(t) = \frac{1}{2}\sum_{j=2}^{\infty} \frac{\sum_{d=1}^{j-1} H\left(\mathbf{W}_k^{(d)}\right)}{\left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right)}. \tag{28}$$

Proof. Having Theorem 3 in mind, we just need to prove that $\sum_{d=1}^{n-1} H\left(\mathbf{W}_n^{(d)}\right) = o(n)$. Let us recall that the random vector $\mathbf{W}_n^{(d)} : \mathcal{R}_n^{(d)} \to \{1,\ldots,n-d\}^d$ describes the split at the root of a tree: precisely that a tree root degree equals d and its subtrees are of sizes $W_{n,1}^{(d)},\ldots,W_{n,d}^{(d)}$. Since the entropy of random variable is upper bounded by the logarithm of the variable image cardinality, we have

$$\sum_{d=1}^{n-1} \mathbb{P}(D_n = d)\, H\left(\mathbf{W}_n^{(D_n)} | \mathbb{P}(D_n = d)\right) \leq \sum_{d=1}^{n-1} \mathbb{P}(D_n = d) \log\left(n^d\right) = \log(n) \sum_{d=1}^{n-1} d\,\mathbb{P}(D_n = d).$$

Observe that $\mathbb{E}(D_n) = \sum_{d=1}^{n-1} d\,\mathbb{P}(D_n = d)$ is the expected value of the general plane recursive tree root degree. From [3] we know that $\mathbb{E}(D_n) = \sqrt{\pi n} + O(1)$, what gives us the desired result. □

Remark 5. Taking closer look at $H\left(\mathbf{W}_n^{(d)}\right)$ we find that

$$H\left(\mathbf{W}_n^{d}\right) = \log\left(n\frac{r_n}{n!}\right) - \sum_{d=1}^{n-1} d \sum_{k=1}^{n-d} q_{n,k}^{(d)} \log\left(\frac{r_k}{k!}\right).$$

This allows us to check numerically that the entropy rate of the unlabeled general plane trees, generated according to the model of plane recursive trees is $h(t) \approx 1.68$.

IV. Concluding Remarks

In this paper we focus on finding entropies of various advanced trees, namely, m-ary search trees, d-ary increasing trees, and general trees. In the course of deriving these entropies we encounter interesting novel recurrences that we show how to solve in their generalities. These recurrences will find ample of applications in analyzing such general trees. For example, as in [13], the next natural question is to find entropy of non-plane d-ary trees and general trees. For arbitrary d, we expect to meet some challenging mathematical problems to find these entropies (see [7]).

We did not address here how to compress optimally these trees. But it is not hard to see that a direct generalization of arithmetic encoding proposed in [13] can be used. More precisely, we need to traverse the tree from left to right and encode the ratio of the number of internal nodes in a subtree to all internal nodes.

Acknowledgment

This work was supported by NSF Center for Science of Information (CSoI) Grant CCF-0939370, by NSF Grant CCF-1524312, and NIH Grant 1U01CA198941-01. Z. Gołębiewski was partially supported by Polish NCN grant 2013/09/B/ST6/02258.

References

[1] M. Adler, M. Mitzenmacher, Towards Compressing Web Graphs, *Data Compression Conference* 2001, pp. 203-212.

[2] D. Aldous and N. Ross, Entropy of Some Models of Sparse Random Graphs With Vertex-Names. *Probability in the Engineering and Informational Sciences*, 2014, 28:145-168.

[3] François Bergeron, Philippe Flajolet, and Bruno Salvy. Varieties of increasing trees. In Jean-Claude Raoult, editor, *CAAP '92, 17th Colloquium on Trees in Algebra and Programming, Rennes, France, February 26-28, 1992, Proceedings*, volume 581 of *Lecture Notes in Computer Science*, pages 24–48. Springer, 1992.

[4] F. Chierichetti, R. Kumar, S. Lattanzi, M. Mitzenmacher, A. Panconesi, and P. Raghavan. On Compressing Social Networks, *Proc. ACM KDD*, 2009.

[5] Hua-Huai Chern and Hsien-Kuei Hwang. Phase changes in random m-ary search trees and generalized quicksort. *Random Struct. Algorithms*, 19(3-4):316–358, 2001.

[6] Yongwook Choi, Wojciech Szpankowski: Compression of Graphical Structures: Fundamental Limits, Algorithms, and Experiments. *IEEE Transactions on Information Theory*, 2012, 58(2):620-638.

[7] J. Cichon, A. Magner, W. Szpankowski, K. Turowski, On symmetries of non-plane trees in a non-uniform model, *ANALCO*, Barcelona, 2017.

[8] Michael Drmota. *Random Trees, An Interplay between Combinatorics and Probability*. Springer-Verlag Wien, 2009.

[9] James Allen Fill and Nevin Kapur. Transfer theorems and asymptotic distributional results for m-ary search trees. *Random Structures & Algorithms*, 26(4):359–391, 2005.

[10] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[11] J. C. Kieffer, E.-H. Yang, W. Szpankowski, Structural complexity of random binary trees. *ISIT 2009* , pp. 635-639.

[12] Donald E. Knuth. *The Art of Computer Programming, Volume 3: (2Nd Ed.) Sorting and Searching*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998.

[13] A. Magner, W. Szpankowski, K. Turowski, Lossless Compression of Binary Trees with Correlated Vertex Names, *ISIT'16*, Barcelona, 2016.

[14] M. Mohri, M. Riley, A. T. Suresh, Automata and graph compression. *ISIT 2015*, pp. 2989-2993.

[15] L. Peshkin, Structure induction by lossless graph compression, *In Proc. of the IEEE Data Compression Conference*, 53–62, 2007.

[16] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*. John Wiley & Sons, Inc., New York, NY, 2001.

[17] J. Sun, E.M. Bollt, and D. Ben-Avraham, Graph compression–save information by exploiting redundancy, *Journal of Statistical Mechanics: Theory and Experiment*, P06001, 2008.

[18] J. Zhang, E.-H. Yang, J. C. Kieffer, A Universal Grammar-Based Code for Lossless Compression of Binary Trees. *IEEE Transactions on Information Theory*, 2014, 60(3):1373-1386.

[19] Daniel (ed.) Zwillinger. *CRC Standard Mathematical Tables and Formulae*. CRC Press, Boca Raton, Florida, 2011.