International Journal of Pure and Applied Mathematics
Volume 81 No. 6 2012, 813-820
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu

ijpam.eu

ANALYSIS OF A SPECIAL TREE STRUCTURE

Ramin Kazemi

Department of Statistics
Imam Khomeini International University
Qazvin, IRAN

Abstract: In this note we discuss on a special tree structure of size n which we call a b-increasing tree for convenience. This family of trees have the total weights $T_n = (n-1)!$. We study the quantity depth of the largest element and show that the corresponding bivariate generating function satisfies a differential equation of order b where b is the maximal bucket size in this model. Also we concentrate on the model, where deterministically weights are attached to the edges according to the definition of actual labels in this tree. We will use a bijection between this model and permutations. Using this bijection we obtain some results on the sum of weighted edges.

AMS Subject Classification: 05C05, 60F99
Key Words: tree structure, b-increasing tree, depth, weighted edge, actual labels, permutation

1. Introduction

In this paper we discuss on a new model of tree structure called b-increasing tree where the nodes are labelled with integers $1, 2, ..., n$. More explicitly, the tree starts with the node labelled by 1 that has a descendant. The next node is filled with labels 1 and 2 that has a descendant which is filled with three elements 1, 2 and 3. In fact, b nodes with capacities $1, 2, ..., b$ respectively, are

Received: July 8, 2012

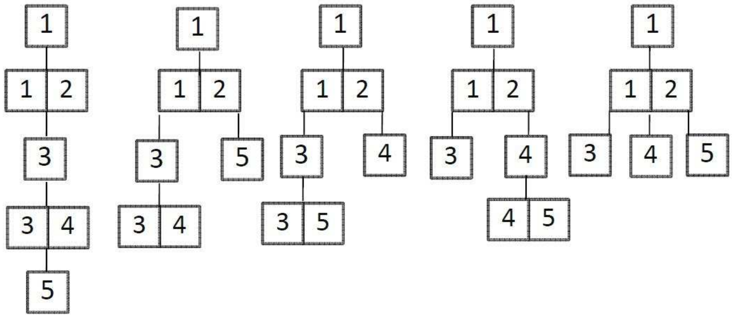

Figure 1: All 2-increasing trees of size 5

connected together with 1 edge such that the elements within these nodes are arranged in increasing order. The node with capacity b has $r \geq 0$ descendants each of them making a subtree such that their nodes are filled similar to the previous structure. The other subtrees are filled just like the first subtree until the label n is inserted in the tree. Figure 1 illustrates all 2-increasing trees of size 5. This model can be considered as a special split tree but labelled with integers 1, 2, ..., n [3].

2. Combinatorial Description of the Model

As mentioned earlier, the nodes are buckets with an integer capacity c and the additional restriction. We always call a node v with capacity $c = b$ saturated and otherwise unsaturated. We define $\lambda_k := \{v \in T;\ c = c(v) = k < b,\ \text{and } v \text{ is a leaf}\}$ and $\mu_k := \{v \in T;\ c = c(v) = b \text{ and } d(v) = k\}$, where $d(v)$ denotes the out-degree of node v. It will be convenient to define the size $|T|$ of a tree T via $|T| = \sum_{v \in \lambda_k \cup \mu_k} c(v)$.

A class $\mathcal{T}$ of a family of b^*-increasing trees can be defined in the following way. A sequence of non-negative numbers $(\alpha_k)_{k \geq 0}$ with $\alpha_0 > 0$ and two sequence of non-negative numbers $\beta_1, \beta_2, ..., \beta_{b-1}$ and $\gamma_1, \gamma_2, ..., \gamma_{b-1}$ are used to define the weight $w(T)$ of any tree T by $w(T) := \Pi_v w(v)$, where $w(v)$ ranges over all nodes of T. We define an appropriate weight $w(v)$ of a node v as follows (regarding to the definition of the model):

$$w(v) = \begin{cases} \alpha_{d(v)}, & v \text{ is saturated} \\ \beta_{c(v)}, & v \text{ is unsaturated (leaf)} \\ \gamma_{c(v)}, & v \text{ is unsaturated (non-leaf).} \end{cases} \tag{1}$$

Thus for saturated nodes the weight is dependent on the out-degree and described by the sequence α_k, whereas for unsaturated nodes the weights are dependent on the capacity and described by the sequences β_k and γ_k.

Let $\mathcal{L}(T)$ denotes the set of different defined labelings of the tree T with distinct integers $\{1, 2, ..., |T|\}$, where $L(T) := |\mathcal{L}(T)|$ denotes its cardinality. Then the family $\mathcal{T}$ consists of all trees T together with their weights $w(T)$ and the set of defined labelings $\mathcal{L}(T)$. For a given degree-weight sequence $(\alpha_k)_{k\geq 0}$ with a degree-weight generating function $\varphi(t) := \sum_{k\geq 0} \alpha_k t^k$ and two bucket-weight sequences $\beta_1, \beta_2, ..., \beta_{b-1}$ and $\gamma_1, \gamma_2, ..., \gamma_{b-1}$, we define now the total weights by $T_n := \sum_{|T|=n} w(T) \cdot L(T)$.

As it is natural in enumeration problems related to labelled structures, we use the exponential generating function $T(z) := \sum_{n=1}^{\infty} T_n \frac{z^n}{n!}$. Thus the exponential generating function of the total weights T_n is characterized by the following differential equation of order b (regard to symbolic method) [2]:

$$T^{(b)}(z) \quad = \quad \left(\prod_{i=1}^{b-1} \gamma_i\right)^{|\mathcal{P}_n|} \varphi\Big(T(z)\Big), \tag{2}$$

with initial conditions $T(0) = 0, T^{(k)}(0) = \Pi_{i=1}^{k-1} \gamma_i \beta_k$ for $1 \leq k \leq b-1$ where $\mathcal{P}_n$ is the set of all trees of size n ($|.|$ denotes the size of sets). This could be done by setting up a recurrence for the total weights T_n:

$$T_n = \left(\prod_{i=1}^{b-1} \gamma_i\right)^{|\mathcal{P}_n|} \sum_{r\geq 0} \alpha_r \sum_{k_1+\cdots+k_r=n-b} T_{k_1}...T_{k_r} \binom{n-b}{k_1, k_2, ..., k_r}.$$

Let p be the probability that element $n+1$ is attracted by node $v \in \lambda_k \cup \mu_k$. If we choose the weights $\gamma_k = 1$, $\beta_k = (k-1)!$ and $\alpha_k = \frac{b^k(b-1)!}{k!}$, we can show that the probability p is $\frac{c(v)}{n}$, which coincides with the stochastic growth rule for b-increasing trees ($\alpha_0 = \beta_b = (b-1)!$).

Thus the exponential generating function $T(z)$ of the total weights T_n of b-increasing trees of size n satisfies the differential equation

$$T^{(b)}(z) = (b-1)! \exp(bT(z)) \tag{3}$$

with initial conditions $T(0) = 0, T^{(k)}(0) = (k-1)!$ for $1 \leq k \leq b-1$. The solution of equation (3) is given by $T(z) = -\log(1-z)$. Hence the total weights of b-increasing trees of size n is given by $T_n = (n-1)!$ for $n \geq 1$.

3. Depth of Element n

Let D_n be the depth of element n, i.e., the number of edges lying on the path from the root node to the node that contains element n, in a b-increasing tree of size n. In order to study D_n for b-increasing trees we consider first the corresponding random variable D_n in a b^*-increasing tree family with arbitrary weight sequences α_k, β_k and γ_k as it is natural in this approach. We introduce the bivariate generating function [7]

$$D(z,v) = \sum_{n\geq 1}\sum_{m\geq 0} P(D_n = m) T_n \frac{z^{n-1}}{(n-1)!} v^m. \tag{4}$$

Theorem 1. *The bivariate generating function $D(z,v)$ satisfy the following differential equation of order b:*

$$\begin{aligned} \frac{d^b}{dz^b} D(z,v) &= \frac{vb!D(z,v)}{(1-z)^b} \\ \frac{d^k}{dz^k} D(z,v)\Big|_{z=0} &= k!, \quad 0 \leq k \leq b-1. \end{aligned} \tag{5}$$

Proof. It is convenient to think of specifically bicolored b^*-increasing trees, where the elements contained in the nodes are colored as follows: element n in a size-n tree is colored red and all elements with a label smaller than n are colored black. We are thus interested in the depth of the red element. We consider now a specific bicolored b^*-increasing tree T of size n and we assume that the first node of T with capacity b has out-degree $r \geq 1$ and the red element is not captured in this node (thus $n > b$). By the same technic of [4]

$$\frac{d^b}{dz^b} D(z,v) = \left(\prod_{i=1}^{b-1} \gamma_i\right)^{|\mathcal{P}_n|} v\varphi'(T(z)) D(z,v),$$

since the elements labelled by 1, 2, ..., b contained in the first $b-1$ nodes and bth node are all colored black.

The initial conditions of the above differential equation are given as follows:

$$\begin{aligned} \frac{d^k}{dz^k} D(z,v)\Big|_{z=0} &= \sum_{m\geq 0} P(D_{k+1} = m) T_{k+1} v^m \\ &= T_{k+1} = T^{(k+1)}(0) = \Pi_{i=1}^{k} \gamma_i \beta_{k+1}. \end{aligned}$$

By using the sequences $\gamma_k = 1$, $\beta_k = (k-1)!$ and $\alpha_k = \frac{b^k(b-1)!}{k!}$ proof is completed. □

The homogeneous differential equation (5) is of Cauchy-Euler-type that has been studied in [8] in the context of eigenvalues of a replacement matrix associated to bucket recursive trees. Thus the Hwang's quasi power theorem [2] follows that the random variable D_n is asymptotically normal distributed with rate of convergence $O((\log n)^{-1/2})$ and

$$\begin{aligned}\mathbf{E}[D_n] &= \frac{1}{\sum_{k=1}^{b}\frac{1}{k}}\log n + O(1),\\ \mathbf{Var}[D_n] &= \frac{\sum_{k=1}^{b}\frac{1}{k^2}}{\left(\sum_{k=1}^{b}\frac{1}{k}\right)^3}\log n + O(1).\end{aligned}$$

4. Weighted Edges

There are several well-known bijections between trees and permutations. These bijections characterize the some parameters in the trees. For example, a rotation bijection characterizes the distribution of the root degree, the number of leaves in a random recursive tree [1, 5]. Here we state a bijection which is appropriate to suitably defined weights on the edges of the our model. With this bijection we are able to relate our parameter in permutations.

Let T denote a bucket tree with maximal bucket-size b. We consider edge-weighted bucket trees, where every edge $e \in E(T)$ of the tree will be weighted deterministically as follows.

Definition 1. Let k_i be the maximum of the labels in every node with capacity $1 \le c \le b$. We call k_i the actual label.

If the edge e_i is adjacent to the nodes with actual labels k_{i-1} and k_i, then we define the weight of the edge e_i as $w_{e_i} := k_i - k_{i-1}$ where node with actual label k_{i-1} is a parent of node with actual label k_i.

In the following, we state a natural bijection which maps inversions in random permutations of $\{1, 2, , ..., n-1\}$ to suitably defined weights on the edges of bucket tree with n nodes. Using this bijection we are able to relate our parameter in permutations similar to [5].

Bijection. Consider a bucket tree T_n of size n and its edge set $E(T_n)$. We enumerate the $n-1$ edges of $E(T_n)$ by $e_{k_2}, e_{k_3}, ..., e_{k_n}$, where e_{k_t}, with $k_t \ge 2$, is defined as the edge $e_{k_t} = (k_{t-1}, k_t)$ connecting nodes with actual labels k_{t-1} and k_t, with $1 \le k_{t-1} \le k_t - 1$. We define the numbers $w_{k_t} := w_{e_{k_t}} = k_t - k_{t-1}$ as the edge-weight of edge e_{k_t} and consider the edge-weight table $(w_{k_2}, w_{k_3}, ..., w_{k_n})$

of T_n. Of course, it holds $1 \leq w_{k_t} \leq k_t - 1$, for $2 \leq k_t \leq n$. If we define $i_{k_t} := w_{k_{t+1}} - 1$, for $1 \leq k_t \leq n-1$, then it holds $0 \leq i_{k_t} \leq k_t - 1$ and the array $(i_{k_1}, i_{k_2}, ..., i_{k_{n-1}})$ corresponds to the inversion table of a permutation σ of $\{1, 2, ..., n\}$, which uniquely determines σ. To construct a bucket tree T of size n from a given permutation σ of $\{1, 2, ..., n\}$ with inversion table $(i_{k_1}, i_{k_2}, ..., i_{k_{n-1}})$ one starts with 1 as the root of T and attaches successively node with actual label k_t, with $2 \leq k_t \leq n$, to node $k_{t-1} = k_t - 1 - i_{k_{t-1}}$, which leads to an edge-weight table $(w_{k_2}, ..., w_{k_n})$ with $w_{k_t} = i_{k_{t-1}} + 1$, for $2 \leq k_t \leq n$ and for a given bucket tree T_n of size n we always have

$$\sum_{2 \leq k_t \leq n} w_{k_t} + 1 = n + \sum_{1 \leq k_t \leq n-1} i_{k_{t-1}}. \tag{6}$$

4.1. Sum of the Weighted Edges

In this section we will study the following parameters in a size-n random our model:

1) $T_n(j)$, which counts the number of edges with weight j,

2) T_n, which counts the sum of the weighted edges.

Let $I(w_e = j)$ be the indicator variable of the event that e has weight j. By the above definition, $T_n(j) := \sum_{e \in E_n} I(w_e = j)$.

Suppose $N_n(j)$ denotes the random variable that counts the number of elements in a random permutation of size n with exactly j inversions. Also let N_n be the number of inversions of a random permutation of size n. We denote by $X \stackrel{d}{=} Y$ the equality in distribution of the random variables X and Y.

Lemma 1. *For $1 \leq j \leq n-1$,*

$$T_n(j) \stackrel{d}{=} N_{n-1}(j-1).$$

Proof. It is immediate consequence of bijection introduced in Section 4. □

Theorem 2. *The random variable $T_n(j)$ satisfies the following distribution equality:*

$$T_n(j) \stackrel{d}{=} \sum_{k=j}^{n-1} B_k,$$

where B_k denotes a Bernoulli distributed random variable $Ber(p)$ and p is the probability of attracting node with actual label k to the already grown tree T. Also, the random variables B_k are mutually independent.

Proof. Let $v_k \stackrel{c}{\to} v_i$ be the event that node with actual label k is attached to node with actual label i (note with actual label k is a child of node with actual label i). We have the following distributional equation of $T_n(j)$:

$$T_n(j) \stackrel{d}{=} \sum_{k=j+1}^{n} I(v_k \stackrel{c}{\to} v_{k-j}).$$

Also these indicators are mutually independent. □

We will denote by $s_{k,t}$ the signless Stirling numbers of the first kind.

Theorem 3. *For $1 \le j \le n-1$,*

$$P(T_n(j) = t) = \frac{\sum_{k=t}^{n-j} \binom{n-k-2}{j-2} \frac{s_{k,t}}{k!}}{\binom{n-1}{j-1}}.$$

Proof. From lemma 1, $P(T_n(j) = t) = P(N_{n-1}(j-1) = t)$ and proof is completed [6]. □

Corollary 4. *a) For $j = o(n)$,*

$$T_n^*(j) = \frac{T_n(j) - (\log n - \log j)}{\sqrt{\log n - \log j}} \stackrel{d}{\to} N(0,1).$$

b) For $n - j = o(n)$, $P(T_n(j) = 0) \to 1$.

Theorem 5. *The distribution of the random variable T_n is given by*

$$T_n \stackrel{d}{=} N_{n-1} + (n-1).$$

The random variables are mutually independent. Also as $n \to \infty$,

$$\frac{T_n - \mathbf{E}(T_n)}{\sqrt{\mathbf{Var}(T_n)}} \stackrel{d}{\to} N(0,1).$$

Proof. The distributional equality is a consequence of (6). Using central limit theorem for the number of inversions N_n proof is completed [9]. □

References

[1] F. Bergeron, P. Flajolet, B. Salvy, Varieties of increasing trees, *Lecture Notes in Computer Science* **581**, (1992), 244-8.

[2] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge (2008).

[3] C. Holmgren, Ph.D thesis, *Split Trees, Cuttings and Explosions*, uu.diva-portal.org/smash/get/diva2:285453/FULLTEXT01 (2010).

[4] M. Kuba, A. Panholzer, A combinatorial approach to the analysis of bucket recursive trees, *Theor. Comp. Sci.,* **411**, No-s: 34-36 (2010), 3253-3273.

[5] M. Kuba, A. Panholzer, On edge-weighted recursive trees and inversions in random permutations, *Discrete. Math*, **308** (2008), 529-540.

[6] M. Kuba, A. Panholzer, On the degree distribution of the nodes in increasing trees, *J. Combin. Theory Ser. A.*, **114**, No. 4 (2007), 597-618.

[7] A. Panholzer, H. Prodinger, The level of nodes in increasing trees revisited, *Random Structures Algorithms*, **31** (2007), 203-226.

[8] H. Mahmoud, R. Smythe, Probabilistic analysis of bucket recursive trees, *Theor. Comp. Sci.*, **144** (1995), 221-249.

[9] V.N. Sachkov, *Probabilistic Methods in Combinatorial Analysis*, Cambridge University Press, New York, NY (1997).