ARTICLE IN PRESS YJCTA:3790

Available online at www.sciencedirect.com
ScienceDirect

Journal of Combinatorial Theory, Series A ••• (••••) •••–•••

Journal of Combinatorial Theory Series A

www.elsevier.com/locate/jcta

Note

A generalized enumeration of labeled trees and reverse Prüfer algorithm

Seunghyun Seo [a], Heesung Shin [b]

[a] Department of Mathematics Education, Cheongju University, Cheongju, Chungbuk 360-764, Republic of Korea
[b] Department of Mathematics, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Republic of Korea

Received 1 February 2006

Abstract

A *leader* of a tree T on $[n]$ is a vertex which has no smaller descendants in T. Gessel and Seo showed that

$$\sum_{T\in\mathcal{T}_n} u^{(\text{\# of leaders in } T)} c^{(\text{degree of 1 in } T)} = uP_{n-1}(1,u,cu),$$

which is a generalization of Cayley's formula, where $\mathcal{T}_n$ is the set of trees on $[n]$ and

$$P_n(a,b,c) = c\prod_{i=1}^{n-1}\bigl(ia + (n-i)b + c\bigr).$$

Using a variation of the Prüfer code which is called a *RP-code*, we give a simple bijective proof of Gessel and Seo's formula.

Keywords: Tree; Prüfer code; Leader; Bijection

1. Introduction

A *tree* on V is an acyclic connected graph with vertex set V. In 1889, Cayley [1] showed that $|\mathcal{T}_n| = n^{n-2}$ $(n \geqslant 1)$, where $\mathcal{T}_n$ is the set of trees on $[n] = \{1, 2, \ldots, n\}$. This result is called *Cayley's formula*. Later, in 1918, Prüfer [4] made the *Prüfer code*, which is a bijection between $\mathcal{T}_n$

E-mail addresses: shyunseo@cju.ac.kr (S. Seo), h.shin@kaist.ac.kr (H. Shin).

0097-3165/$ – see front matter
doi:10.1016/j.jcta.2007.01.010

Please cite this article in press as: S. Seo, H. Shin, A generalized enumeration of labeled trees and reverse Prüfer algorithm, J. Combin. Theory Ser. A (2007), doi:10.1016/j.jcta.2007.01.010

and $[n]^{n-2}$. Assume that edges are directed toward vertex 1 and $\mathrm{indeg}_T(i)$ is the indegree of i in T. By the properties of the Prüfer code, we have

$$\sum_{T\in\mathcal{T}_n}\prod_{i\in[n]} x_i^{\mathrm{indeg}_T(i)} = x_1(x_1+\cdots+x_n)^{n-2},$$

which is a generalization of Cayley's formula.

A tree is called a *rooted tree* if one vertex has been designated as the root. A vertex v in a rooted tree is a *descendant* of u if u lies on the unique path from the root to v. By convention, we consider that unrooted trees are rooted at the smallest vertex. A vertex v of a rooted tree is called a *leader* if v is minimal among its descendants. Note that 'leader' is new terminology for the 'proper vertex' which was introduced by Seo [5].

Recently, Gessel and Seo [3] showed that

$$\sum_{T\in\mathcal{T}_n} u^{\mathrm{lead}(T)} c^{\deg_T(1)} = uP_{n-1}(1,u,cu), \tag{1}$$

where $\mathrm{lead}(T)$ is the number of leaders in T and the homogeneous polynomial $P_n(a,b,c)$ is defined by

$$P_n(a,b,c) = c\prod_{i=1}^{n-1}\bigl(ia+(n-i)b+c\bigr).$$

To prove (1), they used generating functions methods.

In this paper, we prove Eq. (1) by giving an algorithm which produces a code with length $n-1$ from a tree with n vertices.

2. Reverse Prüfer algorithm

The *reverse Prüfer code* (*RP-code*) $\varphi(T)=(\sigma_1,\ldots,\sigma_{n-1})$ of a rooted tree T on $[n]$ is generated by successively selecting the unselected vertex of T having the smallest descendant. (Note that a vertex is a descendant of itself.) If several vertices have the same smallest descendant, we choose the vertex which is closest to the root. To obtain the code from T, each time we select a vertex other than the root (which is always selected first), we record *its parent* σ_i until all the vertices are selected. See Fig. 1 for example. We call this process the *reverse Prüfer algorithm* (*RP-algorithm*).

The inverse of φ may be described as follows. Let $\sigma=(\sigma_1,\ldots,\sigma_{n-1})$ be a sequence of positive integers with $\sigma_i\in[n]$ for all i. We can find the rooted tree T whose code is σ by building trees T_i with i labeled vertices and one unlabeled leaf by reading the code σ from left to right. Before reading the code, we start with the rooted tree T_0 with only one vertex, which is unlabeled. Assume that T_{i-1} is the labeled tree which corresponds to the initial $i-1$ entries $(\sigma_1,\ldots,\sigma_{i-1})$ of σ for $i=1,\ldots,n-1$. We construct T_i as follows. If σ_i does not already occur as a label in T_{i-1}, we assign the label σ_i to the unlabeled leaf of T_{i-1}. If σ_i does occur as a label in T_{i-1}, then we assign the smallest unused label to the unlabeled vertex of T_{i-1}. Then we add a new unlabeled leaf as a child of vertex σ_i. After reading the code σ, we obtain T_{n-1} with $n-1$ labeled vertices and one unlabeled vertex. Finally we obtain T by assigning the remaining unused label in $[n]$ to the unlabeled vertex of T_{n-1}.

Theorem 1. *The map described above is the inverse of the map φ.*

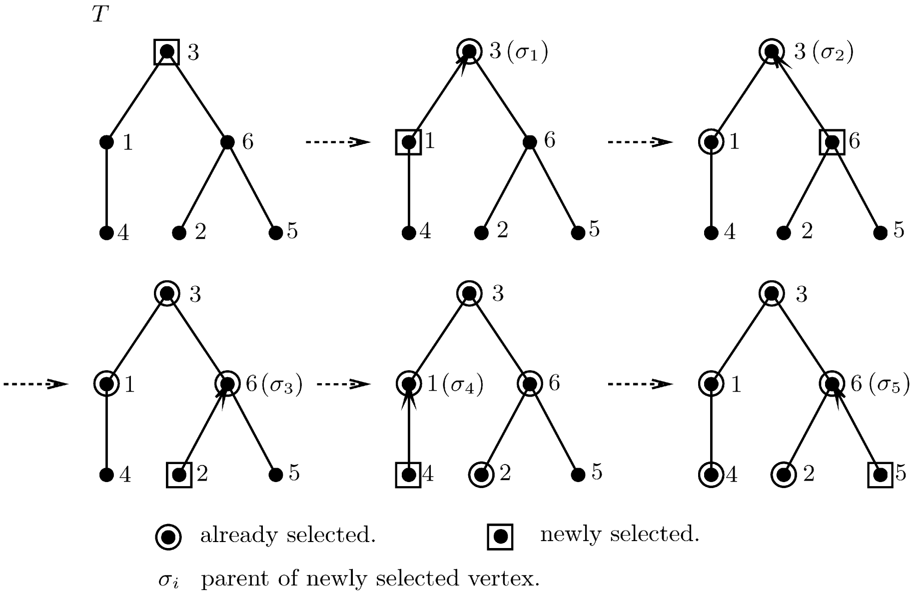

Fig. 1. Rooted tree T to RP-code $\varphi(T) = (3, 3, 6, 1, 6)$.

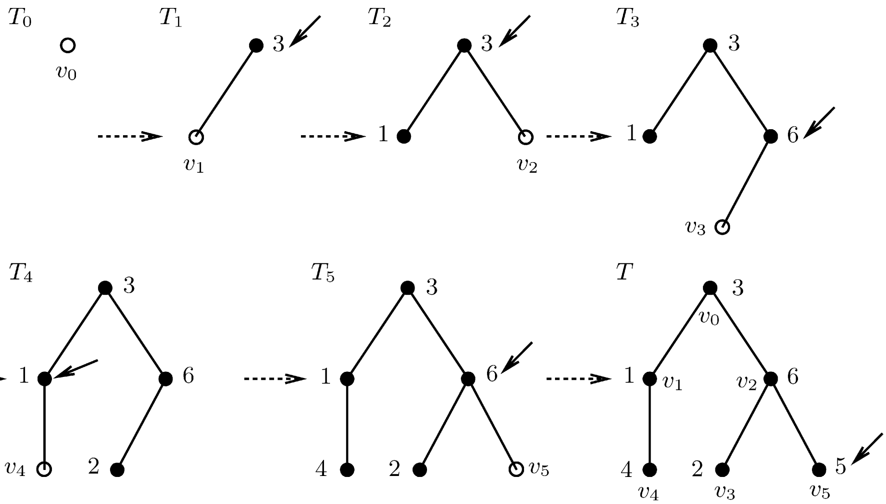

Fig. 2. RP-code $\sigma = (3, 3, 6, 1, 6)$ to tree $\varphi^{-1}(\sigma) = T$.

Proof. Let ψ be the map described above. Consider an arbitrary code C of length $n-1$ and let $T = \psi(C)$. For each $0 \leqslant i \leqslant n-1$, we denote by v_i the unlabeled vertex in T_i, naturally extended to every vertex in T (see Fig. 2). From the label set $[n]$, let l_i be the smallest label not used in T_i. By the definition, l_i is weakly increasing, that is

$$l_0 \leqslant l_1 \leqslant \cdots \leqslant l_{n-1}.$$

If l_i is not used for labeling v_i, then v_i has the descendant v_{i+1} in T_{i+1} and l_{i+1} equals l_i. So, if a vertex v_k is labeled with l_i, then $v_i v_{i+1} \cdots v_k$ becomes a path in T_k. In particular, v_k is a

Please cite this article in press as: S. Seo, H. Shin, A generalized enumeration of labeled trees and reverse Prüfer algorithm, J. Combin. Theory Ser. A (2007), doi:10.1016/j.jcta.2007.01.010

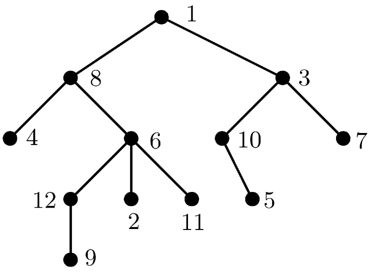

Fig. 3. Example of the tree with root 1.

descendant of v_i in T. Since v_i is a leaf in T_i, v_k is the smallest descendant of v_i in T. Hence l_i is the smallest label among the descendants of v_i in T.

Note that the process of φ is determined by the smallest descendant of each vertex. So φ proceeds according to the order $v_0, v_1, \dots, v_{n-1}$. Hence we have $\varphi(\psi(C)) = \varphi(T) = C$. □

The following theorem justifies the terminology '*reverse*' Prüfer code. Note that, in this paper, the Prüfer code is obtained by successive deletion of the *largest* leaf.

Theorem 2. *If $\varphi(T) = (\sigma_1, \dots, \sigma_{n-1})$ is the RP-code of a rooted tree T, then the Prüfer code of T is $(\sigma_{n-1}, \dots, \sigma_1)$.*

Proof. For each $0 \leqslant i \leqslant n-2$, let T_i' be the labeled tree obtained by deleting the unlabeled vertex from T_{i+1} and let $T_{n-1}' = T$. Since the unlabeled vertex v_i in T_i will be labeled with σ_{i+1} or the smallest unused label in T_i, v_i in T_i' is the maximum leaf and is a child of σ_i. Thus deleting the largest leaf of T_i' and recording its parent yields T_{i-1}' and the entry σ_i. This explains that the RP-code of $T = T_{n-1}'$ equals $(\sigma_{n-1}, \dots, \sigma_1)$. □

The first coordinate of the RP-code of a rooted tree T is always the label of the root of T. In particular, the RP-code of an unrooted tree begins with 1. Cayley's formula is reconfirmed because there are n^{n-2} RP-codes beginning with 1 for trees on $[n]$. Figure 3 shows the tree corresponding to the RP-code $(1, 8, 6, 1, 8, 3, 10, 3, 6, 12, 6)$.

Note that Fleiner [2] found a similar algorithm independently.

3. Statistics of leaders in trees

Now we consider leaders in T during the RP-algorithm. Let $\sigma = (\sigma_1, \dots, \sigma_{n-1})$ be an RP-code. For each $i = 2, \dots, n$, let T_{i-1} be the tree obtained from the subcode $\sigma_1, \dots, \sigma_{i-1}$. Let l be the minimal element in $[n]$ which does not appear in T_{i-1}. To construct T_i from T_{i-1} and σ_i, we should consider the following two cases.

(1) Suppose that σ_i appears in T_{i-1}. Then the unlabeled vertex v in T_{i-1} is labeled by l in T_i. Since the new label l is minimal among the unused labels in T_{i-1}, the vertex v is a leader in T.

(2) Suppose that σ_i does not appear in T_{i-1}. Then the unlabeled vertex v in T_{i-1} is labeled by σ_i in T_i.
   (a) If $\sigma_i = l$, then the vertex v is a leader in T, as in case (1).

Please cite this article in press as: S. Seo, H. Shin, A generalized enumeration of labeled trees and reverse Prüfer algorithm, J. Combin. Theory Ser. A (2007), doi:10.1016/j.jcta.2007.01.010

(b) If $\sigma_i \neq l$, then the vertex v will have a descendant labeled by l. Thus, the vertex v is not leader in T.

So there are exactly i choices of σ_i, cases (1) and (2a), such that the newly labeled vertex v is a leader in T. Because the number of r's ($=\sigma_1$) in a RP-code equals the degree of the root r in T, $\deg_T(1)$ is the number of 1 in the RP-code of an unrooted tree T.

Thus we have the following formula:

$$\begin{aligned}
\sum_{T\in\mathcal{T}_n} u^{\mathrm{lead}(T)} c^{\deg_T(1)} &= cu && \text{by } \sigma_1\ (=1)\\
&\quad \times \big((n-2)+u+cu\big) && \text{by } \sigma_2\\
&\quad \times \big((n-3)+2u+cu\big) && \text{by } \sigma_3\\
&\quad\ \vdots\\
&\quad \times \big(1+(n-2)u+cu\big) && \text{by } \sigma_{n-1}\\
&\quad \times u && \text{by filling the last label}\\
&= cu^2\prod_{i=2}^{n-1}\big((n-i)+(i-1)u+cu\big)\\
&= uP_{n-1}(1,u,cu).
\end{aligned}$$

This completes the bijective proof of Eq. (1).

References

[1] A. Cayley, A theorem on trees, Q. J. Math. 23 (1889) 376–378.

[2] T. Fleiner, On Prüfer codes, Technical report of Egerváry Research Group, TR-2005-16, available at: http://www.cs.elte.hu/egres/www/tr-05-16.html, 2005.

[3] I.M. Gessel, S. Seo, A refinement of Cayley's formula for trees, Electron. J. Combin. 11 (2) (2006) #R27.

[4] H. Prüfer, Neuer Beweis eines Satzes über Permutationen, Arch. Math. (3) 27 (1918) 142–144.

[5] S. Seo, A combinatorial proof of Postnikov's identity and a generalized enumeration of labeled trees, Electron. J. Combin. 11 (2) (2004) #N3.

Please cite this article in press as: S. Seo, H. Shin, A generalized enumeration of labeled trees and reverse Prüfer algorithm, J. Combin. Theory Ser. A (2007), doi:10.1016/j.jcta.2007.01.010