ENUMERATION OF END-LABELED TREES*

ABBE MOWSHOWITZ
Dept. of Computer Science, University of British Columbia, Vancouver, Canada V6T1WE
and
FRANK HARARY
Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48104
And Department of Mathematics, Technion, Israel Institute of Technology, Haifa, Israel

Labeled trees with unlabeled endpoints were counted by Harary, Mowshowitz and Riordan [3]. Moon [5] enumerated connected labeled graphs with unlabeled endpoints. In the present note we examine the complementary problem of counting trees in which only the endpoints are labeled, and in so doing develop a general technique for counting certain classes of partially labeled graphs.

Let $G = (V,X)$ be a graph where $V = \{v_1, v_2, \cdots, v_p\}$ is the set of points, and X its set of lines; see [2]. A *partial labeling* of G is an injection f of $N = \{1, 2, \cdots, n\}$ into V for $n \leqslant p$. A graph G together with a partial labeling f will be called *partially labeled.* Two partially labeled graphs (G, f_1) and (G, f_2) are *identical* if there is an automorphism γ of G such that $f_2(i) = \gamma(f_1(i))$ for $1 \leqslant i \leqslant n$.

A partially labeled tree (T,f) will be called *end-labeled* if $f(N)$ is the set of endpoints of T. Let $t(p)$ and $T(p)$ denote the number of end-labeled trees and end-labeled rooted trees, respectively, having p points.

Theorem 1.

(1)
$$t(p) = B(p-2)$$
and

(2)
$$T(p) = B(p-1),$$
where
$$B(n) = \sum_{k=1}^{n} S(n,k)$$
is a Bell number, i.e., $S(n,k)$ is a Stirling number of the second kind.

Both (1) and (2) follow from the same line of argument so that only (1) will be proved. We will present two derivations of this simple result; the second illustrates a general principle for enumerating partially labeled graphs.

First Proof. Let (T,f) be a p-point end-labeled tree with $V - f(N) = \{v_{n+1}, \cdots, v_p\}$, so that T may be regarded as a labeled tree. Consider the Prufer sequence $(i_1, i_2, \cdots, i_{p-2})$ associated with T (see for example Moon [6] or Harary and Palmer [4]). Each i_j $(1 \leqslant j \leqslant p-2)$ satisfies $n+1 \leqslant i_j \leqslant p$, so that the sequence $(i_1, i_2, \cdots, i_{p-2})$ may be regarded as a distribution of $p-2$ distinct objects into $p-n$ identical cells with no cell empty. The number of such distributions is of course $S(p-2, p-n)$, and hence
$$t(p) = \sum_{n=2}^{p-1} S(p-2, p-n),$$
as asserted.

The second method requires several lemmas. Let U be the set of endpoints of a tree T, and let $\Gamma = \Gamma(T)$ denote its automorphism group. Furthermore, let us define $\Gamma^* = \Gamma^*(T)$ to be the restriction of Γ to U. Then Γ^* is well-defined since U is invariant under any automorphism of T.

*This research was supported in part by grant NRC A-7328 from the National Research Council of Canada and in part by grant 73-2502 from the U.S. Air Force Office of Scientific Research.

Lemma 1. For any tree T, $\Gamma(T)$ is isomorphic to $\Gamma^*(T)$.

Proof. It is clear that the mapping h defined by $\gamma \to \gamma|_U$ for any $\gamma \in \Gamma(T)$ is a homomorphism of Γ onto Γ^*. Now let γ be an arbitrary nontrivial automorphism of T. It is easy to show (see for example Prins [5, p. 17]) that there exist endpoints u and v $(u \neq v)$ such that $\gamma(u) = v$. Hence, h has a trivial kernel.

Lemma 2. Let T be a tree with n endpoints. The number of distinct end-labeled copies of T is $n!/|\Gamma(T)|$.

Proof. Using Lemma 1, this follows from the argument which establishes the analogous result for labeled graphs (see for example Chao [1] or Harary and Palmer [4, p. 4]).

Second Proof of Theorem 1. Let $t^*(p,n)$ and $t(p,n)$ be the number of labeled and end-labeled trees, respectively, having p points n of which end-points. It is well-known that

$$t^*(p,n) = \sum \frac{p!}{|\Gamma(T)|} \tag{3}$$

and by Lemma 2,

$$t(p,n) = \sum \frac{n!}{|\Gamma(T)|}\,, \tag{4}$$

where both summations are over all p-point trees T with n end-points. From (3) we obtain

$$\sum \frac{1}{|\Gamma(T)|} = \frac{1}{p!}\, t^*(p,n)\ ;$$

substituting in (4) gives

$$t(p,n) = \frac{n!}{p!}\, t^*(p,n).$$

Hence,

$$t(p) = \frac{1}{p!} \sum_{n=2}^{p-1} n!\, t^*(p,n),$$

and the result follows from the fact that

$$t^*(p,n) = \frac{p!}{n!}\, S(p-2,\, p-n)$$

(see Moon [4] for several derivations of this formula).

This method of proof illustrates a general counting principle for partially labeled graphs. Let $G = (V,X)$ be a graph which satisfies some given condition A; let S be a property defined on V; and $S(G)$ the subset of V consisting of all points satisfying property S. Denote by $C^*(p,n)$ the number of p-point labeled graphs satisfying condition A for which $|S(G)| = n$, and by $C_S(p)$ the number of p-point S-labeled graphs (only the points in $S(G)$ are labeled) satisfying condition A.

Then the next result is an immediate extension of Theorem 1, in which $S(G)$ plays the role of the endpoints of a tree.

Theorem 2. If $S(G)$ is invariant under every automorphism of G, and for each nontrivial automorphism γ of G, there exist distinct points u and v in $S(G)$ such that $\gamma(u) = v$, then

$$C_S(p) = \frac{1}{p!} \sum n!\, C^*(p,n),$$

where the summation is taken over all n such that $n = |S(G)|$ for some p-point graph G satisfying condition A.

Note that this counting technique is useful only when the number of labeled graphs G satisfying a condition A can be enumerated according to the order of $S(G)$.

REFERENCES

1. C. Y. Chao, "On Groups and Graphs," *Trans. Amer. Math. Soc., 118* (1965), pp. 488–497.
2. F. Harary, *Graph Theory,* Addison-Wesley, Reading, 1969.
3. F. Harary, A. Mowshowitz and J. Riordan, "Labeled trees with Unlabeled End-points," *J. Combinatorial Theory, 6* (1969), pp. 60–64.
4. F. Harary and E.M. Palmer, *Graphical Enumeration,* Academic Press, New York, 1973.
5. J.W. Moon, "Connected Graphs with Unlabeled End-points," *J. Combinatorial Theory 6* (1969), pp. 65–66.

6. J.W. Moon, *Counting Labeled Trees,* Canadian Mathematical Congress, Montreal, 1970.
7. G. Prins, "On the Automorphism Group of a Tree, Doctoral Dissertation, University of Michigan, 1957.

****

[Continued from Page 278.]

$$U(2n+1) = 2T_n + p(n).$$

Secondly, if one places all the partition summands in a line separated by plusses, then one deletes the plus signs at the end of each partition, so that

$$P(n) = U(n) + S(n) - p(n),$$

leading to

$$P(2n) = U(2n) + S(2n) - p(2n) = 2T_n + T_n - p(2n) = 3T_n - n - 1, \quad n \geqslant 1.$$

Equivalently,

$$P(2n+2) = 3T_{n+1} - (n+1) - 1 = \frac{3(n+1)(n+2)}{2} - n - 2$$

$$= \frac{3(n+1)n}{2} + \frac{3(n+1)2 - 2(n+2)}{2}$$

$$= 3T_n + 2n + 1, \quad n \geqslant 0.$$

More easily, we have

$$P(2n+1) = U(2n+1) + S(2n+1) - p(2n+1) = 2T_n + p(2n+1) + T_n - p(2n+1) = 3T_n,$$

which finishes the proof.

We note that the generating function for each sequence given is easily written since the triangular numbers are involved, as

$$\sum_{n=0}^{\infty} P(2n+1)x^n = \frac{3}{(1-x)^3}$$

$$\sum_{n=0}^{\infty} P(2n+2)x^n = \frac{4-x^2}{(1-x)^3}$$

****