COUNTING FORESTS BY DESCENTS AND LEAVES

Ira Gessel

Department of Mathematics
Brandeis University
Waltham, MA 02254-9110
ira@cs.brandeis.edu

Submitted: January 26, 1995; Accepted: June 14, 1995

Dedicated to Dominique Foata

Abstract. A *descent* of a rooted tree with totally ordered vertices is a vertex that is greater than at least one of its children. A *leaf* is a vertex with no children. We show that the number of forests of rooted trees on a given vertex set with $i+1$ leaves and j descents is equal to the number with $j+1$ leaves and i descents. We do this by finding a functional equation for the corresponding exponential generating function that shows that it is symmetric.

Introduction. By a *forest* we mean a forest of rooted labeled trees in which the labels are totally ordered. A *descent* of a tree is a vertex that is greater than at least one of its children. A *leaf* is a vertex with no children.

For a forest F, let $d(F)$ be the number of descents of F and let $l(F)$ be the number of leaves of F. For $n > 0$, let

$$u_n(\alpha, \beta) = \sum_F \alpha^{d(F)} \beta^{l(F)-1},$$

where the sum is over all forests F with vertex set $[n] = \{1, 2, \ldots, n\}$. Since a forest of rooted trees on $[n]$ may be identified with the (unrooted) tree on $\{0, 1, \ldots, n\}$ obtained by joining all the roots of the forest to the new vertex 0, $u_n(\alpha, \beta)$ may also be interpreted in terms of unrooted trees rather than forests of rooted trees.

Our main result is that u_n is symmetric; i.e., $u_n(\alpha, \beta) = u_n(\beta, \alpha)$. More precisely, we shall prove that the exponential generating function

$$U(x; \alpha, \beta) = \sum_{n=1}^{\infty} u_n(\alpha, \beta) \frac{x^n}{n!}$$

satisfies the functional equation

$$1 + U = (1 + \alpha U)(1 + \beta U) e^{x(1-\alpha-\beta-\alpha\beta U)},$$

which implies that U is symmetric in α and β.

We first discuss what is already known about the polynomials $u_n(\alpha,\beta)$. Since there are $(n+1)^{n-1}$ forests of rooted trees with vertex set $[n]$ (see, e.g., Moon [8] for many proofs) we have $u_n(1,1) = (n+1)^{n-1}$. It is also known that the number of forests of rooted trees on $[n]$ with i leaves is $(n!/i!)S(n, n-i+1)$, where $S(n,k)$ is the Stirling number of the second kind. (See Moon [8, p. 20, Theorem 3.5] or Knuth [7; exercise 19, p. 397; solution, p. 585].) Thus

$$u_n(1,\beta) = \sum_{i=0}^{n-1} \frac{n!}{(i+1)!} S(n, n-i)\beta^i.$$

A forest with only one leaf consists of a single "linear" tree, which may be viewed as a permutation, and the descents of the forest are the same as those of the permutation, so $u_n(\alpha, 0) = A_n(\alpha)/\alpha$, where $A_n(\alpha)$ is the nth Eulerian polynomial [4; 9, p. 22].

A tree with no descents is called an *increasing tree* and a forest of increasing trees is called an *increasing forest*. There is a well-known bijection from increasing forests on $[n]$ to permutations of $[n]$ that takes leaves to descents (but we must count an extra descent at the end of the permutation); see, for example, [9, p. 25]. Thus $u_n(0,\beta) = A_n(\beta)/\beta$. Descents of trees seem first to have been considered in [5], where it is shown that $u_n(\alpha, 1) = u_n(1,\alpha)$. This result, together with the other special cases mentioned above, provided the motivation for studying $u_n(\alpha,\beta)$, and suggested that it might be symmetric.

The functional equation. Rather than counting forests directly, we first count some related objects. A *marked forest* is an ordered pair (F, M) where F is a nonempty forest and M is a set of vertices of F containing all of the descents and none of the leaves. We call the vertices in M *marked vertices*. For $n > 0$, let

$$c_n(\beta,\gamma) = \sum_{(F,M)} \beta^{l(F)}\gamma^{|M|}, \tag{1}$$

where the sum is over all marked forests (F, M) with vertex set $[n]$. Since the set of marked vertices in a marked forest consists of all the descents together with an arbitrary subset of the vertices that are neither leaves nor descents, we have

$$\begin{aligned} c_n(\beta,\gamma) &= \sum_F \beta^{l(F)}\gamma^{d(F)}(1+\gamma)^{n-l(F)-d(F)} \\ &= (1+\gamma)^n \sum_F \left(\frac{\beta}{1+\gamma}\right)^{l(F)} \left(\frac{\gamma}{1+\gamma}\right)^{d(F)} \\ &= \beta(1+\gamma)^{n-1} u_n\left(\frac{\gamma}{1+\gamma}, \frac{\beta}{1+\gamma}\right), \end{aligned}$$

and thus

$$u_n(\alpha,\beta) = \beta^{-1}(1-\alpha)^n c_n\left(\frac{\beta}{1-\alpha}, \frac{\alpha}{1-\alpha}\right). \tag{2}$$

Now let

$$C = C(x;\beta,\gamma) = \sum_{n=1}^{\infty} c_n(\beta,\gamma)\frac{x^n}{n!},$$

and let

$$U(x;\alpha,\beta) = \sum_{n=1}^{\infty} u_n(\alpha,\beta)\frac{x^n}{n!}.$$

It follows from (2) that

$$\beta U(x;\alpha,\beta) = C\left(x(1-\alpha); \frac{\beta}{1-\alpha}, \frac{\alpha}{1-\alpha}\right). \tag{3}$$

Next, we describe a decomposition for marked forests that allows us to count them. Let (F, M) be a marked forest and let V_0 be the set of vertices v of F with the property that no (proper) ancestor of v is marked. It is clear that the induced subgraph of F on V_0 is an increasing forest F_0, which we call the *initial forest* of F. Note that every leaf of F_0 is either a leaf or a marked vertex of F, and that if a leaf of F_0 is a marked vertex of F, then its descendents form a marked forest. Thus we can decompose any marked forest into its initial forest together with a (possibly empty) set of marked forests. This decomposition will yield a functional equation for $C(x;\beta,\gamma)$.

To a marked forest (F, M) we assign the weight $\beta^{l(F)}\gamma^{|M|}$, as in (1). Let F_0 be the initial forest of F and let v be a leaf of F_0. If v is a leaf of F then v contributes a factor of β to the weight of F, and if v is a marked vertex of F then v and its descendents contribute to the weight of F a factor of γ times the weight of the marked forest made up of the descendents of v. Now let $A_{n,i}$ be the number of increasing forests on $[n]$ with i leaves. By the properties of exponential generating functions (see, for example, Foata [3], Goulden and Jackson [6, Chapter 3], or Bergeron, Labelle, and Leroux [1, Chapter 5]), it follows that the exponential generating function for marked forests in which the initial forest has n vertices and i leaves is

$$A_{n,i}\frac{x^n}{n!}(\beta+\gamma C)^i.$$

As noted in the introduction, $\sum_i A_{n,i}t^i = A_n(t)$ is the nth Eulerian polynomial. Let

$$A(x;t) = \sum_{n=0}^{\infty} A_n(t)\frac{x^n}{n!},$$

where $A_0(t) = 1$. Then we have

$$1 + C = \sum_{n,i=0}^{\infty} A_{n,i}\frac{x^n}{n!}(\beta+\gamma C)^i = A(x;\beta+\gamma C). \tag{4}$$

It is well known [2, p. 51, equation 14v] that the Eulerian polynomials have the exponential generating function

$$A(x;t) = \sum_{n=0}^{\infty} A_n(t)\frac{x^n}{n!} = \frac{1-t}{1-te^{(1-t)x}}. \tag{5}$$

From (3), (4), and (5) we find that $U = U(x;\alpha,\beta)$ satisfies

$$1+\beta U = \frac{1-\alpha-\beta-\alpha\beta U}{1-\alpha-\beta(1+\alpha U)e^{x(1-\alpha-\beta-\alpha\beta U)}}.$$

Simplifying, we obtain

$$1+U = (1+\alpha U)(1+\beta U)e^{x(1-\alpha-\beta-\alpha\beta U)}. \tag{6}$$

Since (6) is symmetric in α and β, and determines U uniquely, it follows that U is also symmetric.

It is possible to solve (6) by Lagrange inversion to get an explicit formula for the coefficients of U, but this formula seems too complicated to be useful. If we set $\alpha = 1$, (6) reduces to

$$1+\beta U = e^{x\beta(1+U)},$$

which can be solved by Lagrange inversion to give (1). If we set $\beta = 0$ or $\alpha = 0$ in (6), so that we are counting increasing trees by endpoints or permutations by descents, U reduces to a generating function for the Eulerian polynomials. (It differs slightly from (5) since it is normalized differently.)

Tables. Here are the coefficients $u_{n,i,j}$ in $u_n(\alpha,\beta) = \sum_{i,j} u_{n,i,j}\alpha^i\beta^j$ for $n \le 6$.

$n = 1$

$i \backslash j$	0
0	1

$n = 2$

$i \backslash j$	0	1
0	1	1
1	1	0

$n = 3$

$i \backslash j$	0	1	2
0	1	4	1
1	4	5	0
2	1	0	0

$n = 4$

$i \backslash j$	0	1	2	3
0	1	11	11	1
1	11	44	17	0
2	11	17	0	0
3	1	0	0	0

$n = 5$

$i \backslash j$	0	1	2	3	4
0	1	26	66	26	1
1	26	237	288	49	0
2	66	288	146	0	0
3	26	49	0	0	0
4	1	0	0	0	0

$n = 6$

$i \backslash j$	0	1	2	3	4	5
0	1	57	302	302	57	1
1	57	1020	2718	1476	129	0
2	302	2718	3858	922	0	0
3	302	1476	922	0	0	0
4	57	129	0	0	0	0
5	1	0	0	0	0	0

References

1. F. Bergeron, G. Labelle, and P. Leroux, *Théorie des éspèces et combinatoire des structures arborescentes*, Publications du laboratoire de combinatoire et d'informatique mathématique, 19, Université du Québec à Montréal, 1994.
2. L. Comtet, *Advanced Combinatorics*, D. Reidel, Dordrecht, 1974.
3. D. Foata, *La série génératrice exponentielle dans les problèmes d'énumération*, Presses de l'Université de Montréal, 1974.
4. D. Foata and M. P. Schützenberger, *Théorie géométrique des polynômes eulériens*, Lecture Notes in Mathematics 138, Springer-Verlag, Berlin, 1970.
5. I. M. Gessel and K.-Y. Wang, *A bijective approach to the permutational power of a priority queue*, preprint.
6. I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, Wiley, New York, 1983.
7. D. E. Knuth, *Fundamental Algorithms: The Art of Computer Programming, Vol. 1*, Addison-Wesley, Reading, Massachusetts, 1968.
8. J. W. Moon, *Counting Labelled Trees*, Canadian Mathematical Congress, 1970.
9. R. P. Stanley, *Enumerative Combinatorics, Volume 1*, Wadsworth & Brooks/Cole, Monterey, California, 1986.