2 sops only

Scan A3609
A3610

McKeon
Preprint

A3609
A3610

The Expected Number of Symmetries in Locally-Restricted Trees I

Kathleen A. McKeon
Connecticut College
New London, CT

Abstract

Exact formulas are derived for the number of symmetries in several types of unlabeled trees with vertices of restricted degree. The trees are d-trees whose vertices have degree at most d and (1,d)-trees whose vertices have degree 1 or d. These results together with similar results for the number of such trees provide formulas for the expected number of symmetries in these trees.

These trees give rise to significant examples in polymer chemistry. For example, (1,4)-trees represent the alkanes and 4-trees represent the carbon skeletons of alkanes. The expected number of symmetries in such trees is important in the study of collections of molecular species formed during some chemical reaction process.

1. Introduction

The enumeration of trees is an important problem in graph theory with a distinguished history as well as applications to theoretical chemistry. The first major work in this area was performed by Cayley who determined exact formulas for the number of labeled trees [C89], the number of rooted trees [C57] and the number of free trees [C75, C81]. These results were extended and an asymptotic analysis of the numbers was provided by Pólya [P37] and Otter [O48].

Cayley's work [C75] was motivated by the problem of enumerating isomers of alkanes, compounds of carbon and hydrogen atoms which have valencies of 4 and 1 respectively. The alkanes have the general formula C_kH_{2k+2} and can be represented by (1,4)-trees. They are the best documented family of chemical compounds and provide a model for much of chemical theory [GoK73]. Generalizing (1,4)-trees, we have (1,d)-trees, which give rise to other meaningful examples in polymer chemistry. There is a correspondence between (1,d)-trees and d-trees that also has chemical significance. While 4-trees correspond to the carbon skeletons of alkanes [GoK73], d-trees in general correspond to skeleton polymers, i.e., polymer molecules that have been stripped of their reactive end groups [GoT76].

The problem addressed in this paper, the enumeration of symmetries in (1,d)-trees and d-trees for d = 3, 4, is also motivated by chemistry. In the study of collections of molecular species, it is almost always the average of some property over an appropriate class of trees that is required. In computing such an average, it is necessary to assign weights to the various trees in the class so as to reflect the (not usually equal) proportions in which they are formed by the chemical process involved. The proper assignment of weights to the trees often involves the orders of their automorphism groups [GoL75]. Consequently, chemists are interested in the orders of the automorphism groups of large trees of various species such as (1,d)-trees and d-trees.

The tool used to do the counting is a two-variable generating function, an approach that seems to have originated in the work of Etherington [Et38]. For a given class C of trees, let $t(x,y)$ be the generating function in two variables x and y such that the coefficient of $y^m x^n$ is the number of trees T in C of order n in which m is the logarithm base 2 of the order of the automorphism group of T. In $t(x,2)$, the coefficient of x^n is the sum of the orders of the automorphism groups of all such trees.

The technique used to do the counting was developed by Pólya [P37], perfected by Otter [O48] and described as a twenty step algorithm for counting various types of trees by Harary, Robinson and Schwenk [HRS75]. The generating functions $t(x,y)$ and $t(x,2)$ satisfy functional equations from which recurrence relations for their coefficients are determined.

In this paper, the technique is illustrated and results given for (1,3)-trees. Exact formulas are determined for the number of symmetries in both planted and free unlabeled (1,3)-trees. These results together with similar results for the number of (1,3)-trees provide formulas for the expected number of symmetries in these trees. A study of the asymptotic behavior of the number of symmetries in such trees will appear in Part II of this series. A brief sketch of the general approach is provided in the short research announcement [KMPR].

2. Generating Functions

While equations are given for (1,3)-trees only, the method has been used to enumerate symmetries in four types of trees: d-trees and (1,d)-trees for d = 3, 4 [Mc87] and may be applied for higher values of d.

We begin by defining a generating function that counts symmetries in planted trees of the specified type. In general, the results for planted trees provide a means for obtaining the results for free trees. However, removing the root of a planted (1,3)-tree leaves a binary tree that has the same automorphism group as the original planted (1,3)-tree. Thus, when counting symmetries in planted (1,3)-trees we are also counting symmetries in binary trees.

For the planted trees of each type, a two-variable logarithmic generating function is defined as follows:

$$T(x,y) = \sum_{n=1}^{\infty} \sum_{m} T_{m,n} y^m x^n \qquad (2.1)$$

For d-trees, $T_{m,n}$ is the number of planted trees T of order n+1 in which $m = \log_2|\Gamma(T)|$, where $\Gamma(T)$ denotes the automorphism group of T. Every (1,d)-tree, planted or free, has 2 modulo (d-1) vertices. This is taken into account in the definition of T(x,y) for (1,d)-trees. In this example, since every planted (1,3)-tree has an even number of vertices, $T_{m,n}$ is defined to be the number of planted (1,3)-trees on 2n vertices (or binary trees on 2n-1 vertices) with 2^m symmetries .

The values which m may assume in the sum (2.1) depend on both d and the type of tree. Since an automorphism of a rooted tree must leave the root fixed, the order of the automorphism group of a planted (1,3)-tree is of the form 2^m where m is an integer ranging from 0 to n-1.

Note that when y = 1 is substituted in (2.1), T(x,1) counts the number of planted trees of the specified type. Substituting y = 2 in (2.1) results in a one-variable generating function which counts symmetries in planted trees of the specified type. Let S_n be the total number of symmetries in all planted (1,3)-trees on 2n vertices.

$$S_n = \sum_{m} T_{m,n} 2^n \qquad (2.2)$$

and

$$T(x,2) = \sum_{n=1}^{\infty} S_n x^n \qquad (2.3)$$

Similarly, t(x,y) can be defined for free trees. However, we actually only work with t(x,2). Thus, we define

$$t(x,2) = \sum_{n=1}^{\infty} s_n x^n \qquad (2.4)$$

which counts symmetries in free, i.e., unrooted trees. For (1,3)-trees, s_n is the total number of symmetries in all free (1,3)-trees on 2n vertices.

3. Functional Relations

To obtain the exact formulas for the number of symmetries in these trees, functional relations satisfied by $T(x,y)$, $T(x,2)$ and $t(x,2)$ are now derived.

First observe that rooted and planted trees of a specified type can be formed from planted trees of that type. A rooted tree in which the root has degree k is formed by taking a collection of k planted trees and identifying their roots to form the root of the new tree. Adding a new vertex adjacent to the root of this rooted tree results in a planted tree in which the degree of the vertex adjacent to the root is k+1. Based on this observation, relations expressing $T(x,y)$ in terms of $T(x,y)$, $T(x^2,y^2)$, and $T(x^3,y^3)$ are derived.

Theorem 3.1 The generating functions $T(x,y)$ and $T(x,2)$ which count symmetries in planted (1,3)-trees satisfy

$$T(x,y) = x + \frac{1}{2} T(x,y)^2 + (y - \frac{1}{2}) T(x^2,y^2) \qquad (3.1)$$

and

$$T(x,2) = x + \frac{1}{2} T(x,2)^2 + \frac{3}{2} T(x^2,4) \qquad (3.2)$$

Proof: The vertex adjacent to the root of a planted (1,3)-tree has degree 1 or 3. The term x counts the symmetries in a planted K_2, the only planted (1,3)-tree in which the vertex adjacent to the root has degree 1.

To count symmetries in those trees in which the vertex adjacent to the root has degree 3, two cases must be considered. Suppose T is the planted (1,3)-tree formed from the planted (1,3)-trees T_1 and T_2 in the manner described above. If $T_1 \neq T_2$, then we have $|\Gamma(T)| = |\Gamma(T_1)|\,|\Gamma(T_2)|$. Then $1/2(T(x,y)^2 - T(x^2,y^2))$ counts symmetries in this case. If $T_1 = T_2$, then we have $|\Gamma(T)| = 2|\Gamma(T_1)|^2$ since the two branches T_1 and T_2 can be permuted. This case is handled by $yT(x^2,y^2)$ with the factor of y accounting for the additional factor of 2 in the group order. Now (3.2) is obtained by substituting $y = 2$ in (3.1). □

Using the following lemma which relates the order of the automorphism group of a free tree to the orders of the automorphism groups of the vertex and edge-rooted versions of the tree, $t(x,2)$ is expressed in terms of $T(x,2)$, $T(x^2,4)$ and $T(x^3,8)$.

Lemma 3.2 For any tree T,

$$|\Gamma(T)| = \sum_{T_1}|\Gamma(T_1)| - \sum_{T_2}|\Gamma(T_2)| + |\Gamma(T_3)| \tag{3.3}$$

where the first sum is taken over all different vertex-rooted versions T_1 of T and the second sum is taken over all different edge-rooted versions T_2 of T. If T has a symmetry edge, an edge whose vertices are interchanged by some automorphism of T, then $T_3 = T$. If T does not have a symmetry edge, then T_3 is the empty graph and $|\Gamma(T_3)| = 0$.

Proof: This lemma is a variation of a lemma due to Otter [HP73]. As in the proof Otter's lemma, the vertex and edge-rooted versions of T can be paired such that the paired vertex and edge-rooted versions of T have the same automorphism group. Recall that an automorphism of a graph must leave the root fixed while an automorphism of an edge-rooted graph must leave the vertices of the root edge fixed. For each vertex v that is not in the center of T, match the version of T that is rooted at v with the edge-rooted version that is rooted at the first edge on the path from v to the center of T.

If T has a symmetry edge, the center of T consists of two vertices, say u and v. Since the edge uv is a symmetry edge of T, rooting T at u is equivalent to rooting T at v. Hence if the version of T that is rooted at the vertex v is paired with the version of T that is rooted at the edge uv, then the difference of the two sums in (3.3) is 0 and $T_3 = T$. Thus, (3.3) holds in this case.

If T does not have a symmetry edge two cases must be considered. If the center of T consists of two vertices u and v, match the version of T that is rooted at v with the version of T that is rooted at the edge uv. In this case and the case that the center of T consists of just one vertex u, there is one vertex-rooted version of T that cannot be paired with an edge-rooted version. This is the tree that results from rooting T at the vertex u which is in the center of T. Since T does not have a symmetry edge, the vertices in the center of T are all fixed points of the automorphisms of the unrooted tree T. Hence this extra vertex-rooted version of T has the same automorphism group as T and (3.3) holds in this case also. □

This lemma can be extended to a statement about the generating functions that count symmetries by multiplying (3.3) by x^n and summing over all trees of the appropriate order. Summing the result over all $n \geq 1$ gives $t(x,2)$ on the left side. The first sum on the right side gives the series that counts symmetries in rooted trees and the second sum gives the series that counts symmetries in edge-rooted trees while $|\Gamma(T_3)|x^n$ sums to the series that counts symmetries in trees with a symmetry edge. For (1,3)-trees we have the following functional relation.

Theorem 3.3 The generating function t(x,2) for symmetries in free (1,3)-trees is given by

$$t(x,2) = \frac{1}{2x}T(x,2)^2 - \frac{1}{3x}T(x,2)^3 + \frac{3}{2x}T(x^2,4) + \frac{13}{3x}T(x^3,8). \qquad (3.4)$$

Proof: First we determine an expression for the series that counts symmetries in rooted (1,3)-trees. As previously described, this expression can be found by using planted (1,3)-trees to build rooted (1,3)-trees. The series for rooted (1,3)-trees is equal to

$$T(x,2) + \frac{3!}{x}T(x^3,8) + \left[\frac{2}{x}T(x^2,4)\,T(x,2) - T(x^3,8)\right]$$
$$+ \left[\frac{1}{3!\,x}(T(x,2)^3 - 3\,T(x^2,4)\,T(x,2) + 2\,T(x^3,8))\right] \qquad (3.5)$$

Symmetries in rooted (1,3)-trees in which the root has degree 1, i.e., in planted (1,3)-trees are counted by T(x,2). To count symmetries in rooted (1,3)-trees in which the root has degree 3, three cases must be considered. Suppose T is the rooted tree formed from the planted (1,3)-trees T_1, T_2 and T_3. The second term of (3.5) counts symmetries in the case that all three trees are the same. The case that exactly two of the three trees are the same and the case that all three are different are handled by the first and second bracketed terms of (3.5) respectively.

A tree rooted at an edge can be formed by identifying the edges incident to the roots of two planted trees. That edge is the root edge of the edge-rooted tree. When the two trees which are combined are the same, that edge is a symmetry edge. Thus,

$$\frac{1}{2x}(T(x,2)^2 + T(x^2,4)) \qquad (3.6)$$

counts symmetries in edge-rooted (1,3)-trees and

$$\frac{2}{x}T(x^2,4) \qquad (3.7)$$

counts symmetries in (1,3)-trees that have a symmetry edge. Combining (3.5), (3.6) and (3.7) as in lemma 3.2 and using the functional relation (3.2) to simplify gives equation (3.4). □

4. Recurrence Relations

From the functional equation (3.1) recurrence relations for $T_{m,n}$, the coefficient of $y^m x^n$ in $T(x,y)$ can now be determined. Note that throughout this section the subscripts on the variables are always non-negative integers. Otherwise one can assume the value of the variable is zero.

The only planted (1,3)-tree with the identity group as its automorphism group occurs when $n = 1$. Otherwise, there is at least one pair of end vertices that can be permuted. Thus, $T_{0,1} = 1$ and $T_{0,n} = 0$ if $n \geq 2$. For $n \leq 2$ and $1 \leq m \leq n-1$, $T_{m,n}$, the number of planted (1,3)-trees on 2n vertices (binary trees on 2n-1 vertices) with 2^m symmetries is as follows:

$$T_{m,n} = T_{(m-1)/2,n/2} - \frac{1}{2} T_{m/2,n/2} + \begin{cases} 0, & \text{if } m = n-1 \\ \frac{1}{2} \sum_{k=1}^{n-1} \sum_{i} T_{i,k} \, T_{m-i,n-k}, & \text{if } m \neq n-1 \end{cases} \qquad (4.1)$$

Recall that S_n is the coefficient of x^n in $T(x,2)$ and let B_n and C_n be the coefficients of x^{2n} and x^{3n} in $T(x^2,4)$ and $T(x^3,8)$ respectively. Then as a consequence of equation (3.4), s_n, the coefficient of x^n in $t(x,2)$, can be expressed in terms of S_n, B_n and C_n. For $n \geq 2$,

$$s_n = \frac{1}{2} \sum_{k=1}^{n} S_k S_{n-k+1} - \frac{1}{3} \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} S_i S_j S_{n-i-j+1} + \frac{3}{2} B_{(n+1)/2} + \frac{13}{3} C_{(n+1)/3} \qquad (4.2)$$

Let T_n be the number of planted (1,3)-trees on 2n vertices (binary trees on 2n-1 vertices) and let t_n be the number of free (1,3)-trees on 2n vertices. Then equations for T_n and t_n [BaKP81] can be combined with (4.1) and (4.2) respectively to give formulas for the expected number of symmetries in these trees. That is, the expected number of symmetries in a planted (1,3)-tree on 2n vertices (binary tree on 2n-1 vertices) is S_n/T_n and the expected number of symmetries in a free (1,3)-tree on 2n vertices is s_n/t_n.

5. Numerical Results

Values of $T_{m,n}$, S_n and s_n were computed using the CDC Cyber 750 in the Computer Laboratory at Michigan State University. The computation of these numbers was limited by the available accuracy and storage restrictions. Another limiting factor was the time required to compute the values using the recurrence relations. In the case of (1,3)-trees, the Fortran programs used to compute S_n for $n \leq 50$ took 52 seconds while an additional 300 seconds were required to compute S_{68}. Table 1 contains values of $T_{m,n}$ for n-6 to 13. Table 2 contains values of S_n and s_n for n-6 to 25.

Table 1. Coefficients of T(x,y) for Planted (1,3)-trees

n	m	$T_{m,n}$	n	m	$T_{m,n}$
6	1	1	11	1	1
	2	2		2	16
	3	2		3	50
	4	1		4	58
				5	54
7	1	1		6	17
	2	4		7	8
	3	3		8	3
	4	3			
8	1	1	12	1	1
	2	6		2	20
	3	7		3	85
	4	7		4	119
	5	1		5	126
	6	0		6	61
	7	1		7	27
				8	9
9	1	1		9	2
	2	9		10	1
	3	14			
	4	14			
	5	6			
	6	1			
	7	1			

No

Ulze new new Ulze

Table 2. Coefficients of T(x,2) and t(x,2) for (1,3)-trees

A3609 A3610

n	S_n	s_n
6	42	24
7	90	168
8	354	240
9	758	608
10	2290	920
11	6002	5680
12	18410	6104
13	51310	18416
14	154106	43008
15	449322	148152
16	1384962	325608
17	4089174	980840
18	12475362	2421096
19	37746786	7336488
20	116037642	19769312
21	355367310	58192608
22	1097869386	164776248
23	3393063162	502085760
24	10546081122	1427051544
25	32810171382	4261678656

Values of T_n and t_n and the ratios S_n/T_n and s_n/t_n were computed using the Macintosh II microcomputer. Table 3 contains values of S_n/T_n, the expected number of symmetries in a planted (1,3)-tree on 2n vertices (binary tree on 2n-1 vertices), and s_n/t_n, the expected number of symmetries in a free (1,3)-tree on 2n vertices, for n=6 to 25.

Symmetries # in planted (1,3) trees on #2n# nodes.

etc

Table 3. **Expected Number of Symmetries in Planted and Free (1,3)-trees**

n	Planted	Free
6	7.000	12.000
7	8.182	42.000
8	15.391	40.000
9	16.478	55.273
10	23.367	51.111
11	28.995	153.514
12	40.820	92.485
13	52.197	136.415
14	70.723	162.294
15	92.644	268.391
16	127.002	287.640
17	166.017	406.988
18	222.731	474.911
19	295.100	665.743
20	395.295	829.041
21	525.569	1114.097
22	702.244	1435.383
23	935.721	1973.833
24	1250.072	2522.478
25	1667.160	3367.375

References

[BaKP81] C. K. Bailey, J. W. Kennedy and E. M. Palmer, Points by degree and orbit size in chemical trees I, Theory and Applications of Graphs (G. Chartrand et al, eds.) Wiley, New York (1981) 27-43.

[BaKP83] C. K. Bailey, J. W. Kennedy and E. M. Palmer, Points by degree and orbit size in chemical trees II, Discrete Appl. Math., **5** (1983) 157-164.

[C57] A. Cayley, On the theory of the analytical forms called trees, Philos. Mag. (4) **13** (1857) 172-176 = Math. Papers, Vol. 3, 242-246.

[C75] A. Cayley, On the analytical forms called trees, with application to the theory of chemical combinations, Rep. Brit. Assoc. Advance. Sci. **45** (1875) 257-305 = Math. Papers, Vol. 9, 427-460.

[C81] A. Cayley, On the analytical forms called trees, Amer. J. Math. **4** (1881) 266-268 = Math. Papers, Vol. 11, 365-367.

[C89] A. Cayley, A theorem on trees, Quart. J. Pure Appl. Math. **23** (1889) 376-378 = Math. Papers, Vol. 13, 26-28.

[Et38] I. M. H. Etherington, On non-associative combinations, Proc. Roy. Soc. Edinburgh **59** (1938/39) 153-162.

[GoK73] M. Gordon and J. W. Kennedy, The graph-like state of matter, Part 2, LGCI schemes for the thermodynamics of alkanes and the theory of inductive inference, J. Chem. Soc., Farady II **69** (1973) 484-504.

[GoL75] M. Gordon and C. G. Leonis, Combinatorial shortcuts to statistical weights and enumeration of chemical isomers, Proc. 5th British Combinatorial Conf. (1975) 231-238.

[GoT76] M. Gordon and W. B. Temple, The graph-like state of matter and polymer science, Chapter 10 in Chemical Applications of Graph Theory (A. T. Balaban, ed.), Academic, London (1976).

[HP73] F. Harary and E. M. Palmer, Graphical Enumeration, Academic, New York (1973).

[HRS75] F. Harary, R. W. Robinson and A. J. Schwenk, Twenty-step algorithm for determining the asymptotic number of trees of various species, J. Austral. Math. Soc. Ser. A **20** (1975) 483-503.

[KMPR] J. W. Kennedy, K. A. McKeon, E. M. Palmer, R. W. Robinson, Asymptotic number of symmetries in locally-restricted trees, Discrete Appl. Math., to appear.

[Mc87] K. A. McKeon, Enumeration of symmetries in locally-restricted trees, Dissertation, Michigan State University, (1987).

[O48] R. Otter, The number of trees, Ann. of Math. **49** (1948) 583-599.
[P37] G. Pólya, Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen, Acta Math., **68** (1937) 145-254.