Journal of Integer Sequences, Vol. 8 (2005), Article 05.1.8

A Combinatorial Interpretation for a Super-Catalan Recurrence

David Callan
Department of Statistics
University of Wisconsin-Madison
Medical Science Center
1300 University Ave.
Madison, WI 53706-1532
USA
callan@stat.wisc.edu

Abstract

Nicholas Pippenger and Kristin Schleich have recently given a combinatorial interpretation for the second-order super-Catalan numbers $(u_n)_{n\geq 0} = (3, 2, 3, 6, 14, 36, \ldots)$: they count "aligned cubic trees" on n interior vertices. Here we give a combinatorial interpretation of the recurrence $u_n = \sum_{k=0}^{n/2-1} \binom{n-2}{2k} 2^{n-2-2k} u_k$: it counts these trees by number of deep interior vertices where "deep interior" means "neither a leaf nor adjacent to a leaf".

1 Introduction

For fixed integer $m \geq 1$, the numbers

$$u_n = \frac{\binom{2m}{m}\binom{2n}{n}}{2\binom{m+n}{m}}$$

satisfy the recurrence relation

$$u_n = \sum_{k\geq 0} 2^{n-m-2k} \binom{n-m}{2k} u_k \qquad (1)$$

and hence are integers except when $m = n = 0$ [1]. We will call them *super-Catalan numbers of order m* (although other numbers go by this name too). For $m = 0$ and $n \geq 1$, they are the odd central binomial coefficients $(1, 3, 10, 35, \ldots)$ (Sloane's A001700, and count lattice paths of n upsteps and $n-1$ downsteps. For $m = 1$ and $n \geq 0$, they are the familiar Catalan numbers $(1, 1, 2, 5, 14, 42, \ldots)$ (Sloane's A000108) with numerous combinatorial interpretations [2, Ex 6.19].

Three recent papers give combinatorial interpretations for $m = 2$, in terms of (i) pairs of Dyck paths [3], (ii) "blossom trees" [4], and (iii) "aligned cubic trees" [5]. In this case the sequence is $(3, 2, 3, 6, 14, 36, \ldots)$ and is Sloane's A007054. The object of this note is to establish a combinatorial interpretation of the recurrence (1) for $m = 2$: it counts the just mentioned aligned cubic trees by number of vertices that are neither a leaf nor adjacent to a leaf.

For $m = 1$, (1) is known as Touchard's identity, and in Section 2 we recall combinatorial interpretations of the recurrence for the cases $m = 0, 1$. In Section 3 we define aligned cubic trees and establish notation. In Section 4 we introduce configurations counted by the right side of (1) for $m = 2$, and in Section 5 we exhibit a bijection from them to size-n aligned cubic trees.

2 Recurrence for $m = 0, 1$

For $m = 0$, (1) counts lattice paths of n upsteps (U) and $n - 1$ downsteps (D) by number k of DUUs where, for example, $DDUUUDDUU$ has 2 DUUs. It also counts paths of n Us and n Ds that start up by number $2k$ (necessarily even) of inclined twinsteps (UU or DD) at *odd* locations. For example, $U_1 U_2 U_3 D_4 D_5 D$ has four inclined twinsteps, at locations $1, 2, 4$ and 5, but only the first and last are at odd locations.

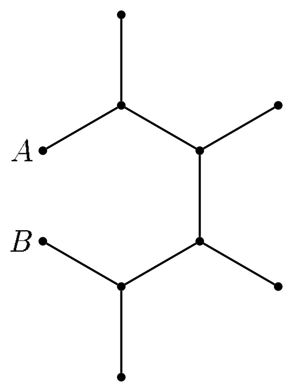
For $m = 1$, (1) counts Dyck n-paths (paths of n Us and n Ds that never dip below ground level) by number k of DUUs. It also counts them by number $2k$ of inclined twinsteps at *even* locations. See [6], for example, for relevant bijections. Recall the standard "walk-around" bijection from full binary trees on $2n$ edges to Dyck n-paths: a worm crawls counterclockwise around the tree starting just west of the root and when an edge is traversed for the first time, records an upstep if the edge is west-leaning and a downstep if it is east-leaning. This bijection carries deep interior vertices to DUUs where deep interior means "neither a leaf nor adjacent to a leaf". Hence the recurrence also counts full binary trees on $2n$ edges by number of deep interior vertices. The interpretation for $m = 2$ below is analogous to this one.

3 Aligned cubic trees

It is well known that there are C_n (Catalan number) full binary trees on $2n$ edges. Considered as a graph, the root is the only vertex of degree 2 when $n \geq 1$. To remedy this, add a vertical planting edge to the root and transfer the root to the new vertex. Now every vertex has degree 1 or 3 and, throughout this paper, we will refer to a vertex of degree 3 as a *node* and of degree 1 as a *leaf*. Thus our planted tree of $2n + 1$ edges has n nodes. Leave the root edge pointing South and align the other edges so that all three angles at each node are

120°, lengthening edges as needed to avoid self intersections. Rotate these objects through multiples of 60° to get all "rooted aligned cubic" trees on n nodes ($6C_n$ of them, since the edge from the root is no longer restricted to point South but may point in any of 6 directions). Now erase the root on each to get all (unrooted) aligned cubic trees on n nodes ($\frac{6}{n+2}C_n$ of them since for each, a root could be placed on any of its $n+2$ leaves). Thus two drawings of an aligned cubic tree are equivalent if they differ only by translation and length of edges. This is interpretation (iii) of the second order super-Catalan numbers mentioned in the Introduction. For short, we will refer to an aligned cubic tree on n nodes simply as an n-ctree (c for cubic).

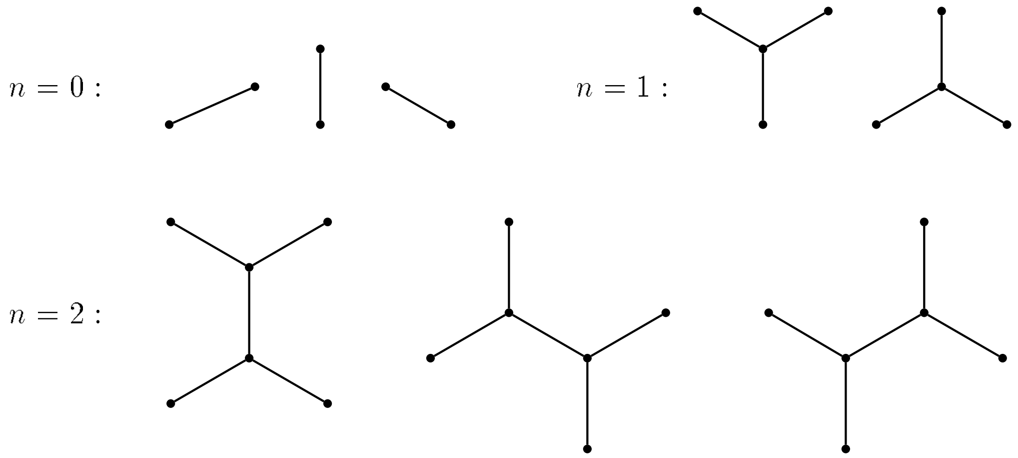
More concretely, a rooted n-ctree can be coded as a pair (r, u) with r an integer mod 6 and u a nonnegative integer sequence of length $n+2$. The integer r gives the angle (in multiples of 60°) from the direction South counterclockwise to the direction of the edge from the root. The sequence $u = (u_i)_{i=1}^{n+2}$ measures distance between successive leaves: traverse the tree in preorder (a worm crawls counterclockwise around the tree starting at a point just right of the root when looking from the root along the root edge). Then $u_i = v_i - 2$ where v_i (≥ 2) is the number of edges traversed between the ith leaf and the next one. For example, the sketched 3-ctree when rooted at A is coded by $(2, (3, 0, 1, 1, 1, 0))$ and when rooted at B is coded by $(1, (0, 1, 1, 1, 0, 3))$.



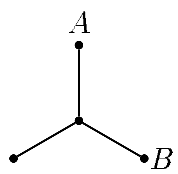
In general, if a ctree rooted at a given leaf is coded by $(r, (u_1, u_2, \ldots, u_{n+1}, u_{n+2}))$ then, when rooted at the next leaf in preorder, it is coded by $((2 + r - u_1) \bmod 6, (u_2, u_3, \ldots, u_{n+2}, u_1))$. Repeating this $n+2$ times all told rotates u back to itself and gives "r" $= 2n+4+r-\sum_{i=1}^{n+2} u_i$ mod 6. Since $\sum_{i=1}^{n+2} u_i$ is necessarily $= 2n-2$, we are, as expected, back to the original coding sequence.

An ordinary (planted) full binary tree is coded by $(0, u)$ and so there are C_n coding sequences of length $n+2$. They can be generated as follows. A ctree can be built up by successively adding two edges to a leaf to turn it into a node. The effect this has on the coding sequence is to take two consecutive entries u_i, u_{i+1} (subscripts modulo $n+2$) and replace them by the three entries $u_i + 1$, 0, $u_{i+1} + 1$. The 1-ctree has coding sequence (0,0,0). The 2-ctree coding sequences are (1,0,1,0) and (0,1,0,1), and so on. Reversing this procedure gives a fast computational method to check if a given u is a coding sequence or not. For example, successively pruning the first 0, 11210230 $\rightarrow$ 1120130 $\rightarrow$ 111030 $\rightarrow$ 11020 $\rightarrow$ 1010 $\rightarrow$ 000 is indeed a coding sequence. However, we will work with the graphical depiction of a ctree.

The n-ctrees for $n = 0, 1, 2$ are shown below. Note that since edges have a fixed non-horizontal direction, we can distinguish a top and bottom vertex for each edge.



It is convenient to introduce some further terminology. Recall a node is a vertex of degree 3. A node is *hidden, exposed, naked* or *stark naked* according as its 3 neighbors include $0, 1, 2$ or 3 leaves. Thus a deep interior vertex is just a hidden node. A 0-ctree has no nodes. Only a 1-ctree has a stark naked node, and hidden nodes don't occur until $n \geq 4$. For $n \geq 2$, an n-ctree containing k hidden nodes has $k + 2$ naked nodes and hence $n - 2k - 2$ exposed nodes. The terms right and left can be ambiguous: we always use right and left relative to travel *from* a specified vertex or edge. Thus vertex B below is left (not right!) travelling from vertex A.

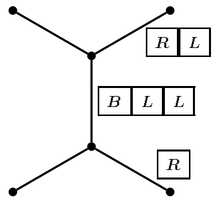


Each n-ctree has a unique center, either an edge or a node, defined as follows. For $n = 0$, it is the (unique) edge in the ctree. For $n = 1$, it is the (unique) node in the ctree. For $n \geq 2$, delete the leaves (and incident edges) adjacent to each naked node, thereby reducing the number of nodes by at least 2. Repeat until the $n = 0$ or 1 definition applies. Equivalently, define the depth of a node in a ctree to be the length (number of edges) in the shortest path from the node to a naked node. Then there are either one or two nodes of maximal depth; if one, it is the center and if two, they are adjacent and the edge joining them is the center.

4 $(n\,,k)$-Configurations

There are $\binom{n-2}{2k}2^{n-2-2k}u_k$ configurations formed in the following way. Start with a k-ctree—u_k choices. Break a strip of $n-2-2k$ squares—$\underbrace{\square\square\square\square\ldots\square\square}_{n-2-2k}$—into $2k+1$ (possibly empty) substrips, one for each of the $2k+1$ edges in the k-ctree—$\binom{(n-2-2k)+(2k+1)-1}{n-2-2k} = \binom{n-2}{2k}$ choices. Mark each square L (= left) or R (= right)—2^{n-2-2k} choices.

Actually, this is not quite what we want. Perform one little tweaking: if there is a center edge and it has an *odd-length* strip of squares, mark the first square T (= top) or B (= bottom) instead of L or R. So a configuration might look as follows ($n = 12$, $k = 2$, empty strips not shown).



5 Bijection

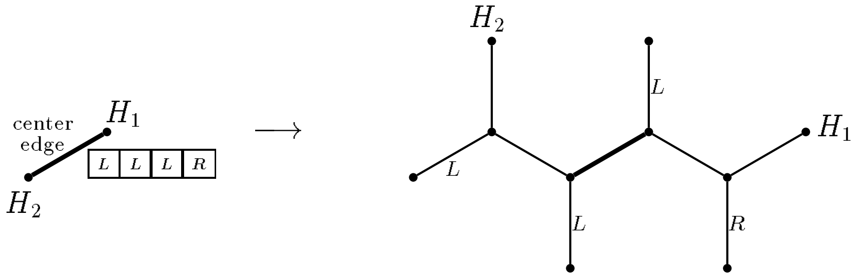
Here is a bijection from (n, k)-configurations to n-ctrees with k hidden nodes. The bijection produces the correspondences in the following table.

(n,k)-configuration	**n-ctree**
leaf	naked node
node	hidden node
square	exposed node

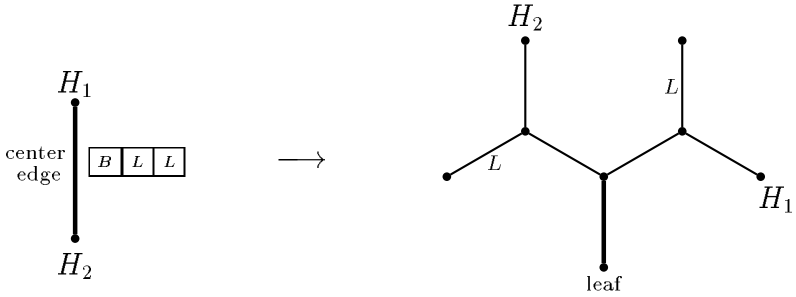
Roughly speaking, work outward from the center, turning a strip of j labeled squares on an edge AB into j exposed nodes lying between A and B.

First, for the center edge (if there is one), the procedure depends on whether it has an even or odd number of squares.

Case Even Here, the center edge becomes an edge joining two exposed nodes as shown: the labels again indicate the L/R status (travelling from the center edge) of the leaves associated with the exposed nodes. The labels are applied from the bottom vertex subtree (H_2) to the top one (H_1).

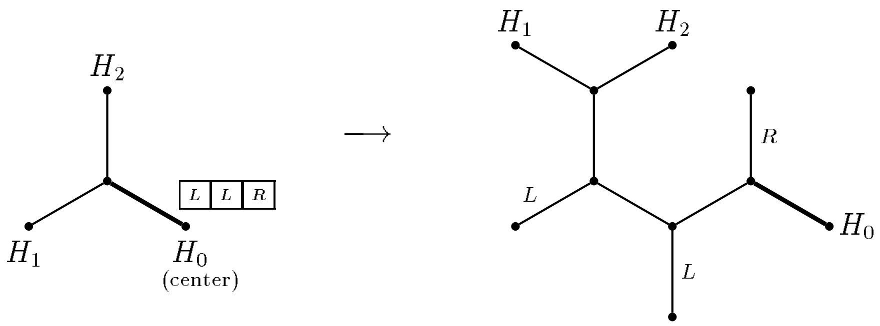


Case Odd Here, the center edge becomes a leaf edge. The first square indicates whether the top or bottom vertex becomes a leaf. Construct equal numbers of exposed nodes on each side of the non-leaf (here, top) vertex, using the L/R designations to determine the leaves (L/R relative to travel from the leaf and running, say, from the left branch to the right).



Decide the placement of the two subtrees H_1, H_2, say the H originally sitting at the vertex which is now a leaf goes at the end of the left branch from the leaf.

Next, for a non-center edge with $i \geq 0$ squares, identify its endpoint closest to the center, let H_0, H_1, H_2 be the subtrees as illustrated (H_0 containing the center), and insert i exposed nodes as shown. The labels L, L, R apply in order from H_1-H_2 to H_0 and indicate the L/R status (travelling from the center) of the leaves associated with the exposed nodes.



Finally, turn the original leaves into naked nodes by adding two edges apiece.

We leave to the reader to verify that the resulting ctree has n nodes of which k are hidden, and that the original configuration can be uniquely recovered from this n-ctree by reversing the above procedure, working in from the leaves.

References

[1] Ira Gessel, Super ballot numbers, *J. Symbolic Computation* **14** (1992), 179–194.

[2] Richard P. Stanley, *Enumerative Combinatorics Vol. 2*, Cambridge University Press, 1999. Exercise 6.19 and related material on Catalan numbers are available online at http://www-math.mit.edu/~rstan/ec/ .

[3] Ira M. Gessel and Guoce Xin, A combinatorial interpretation of the numbers $6(2n)!/(n!(n+2)!)$, math.CO/0401300, 2004, 11pp.

[4] Gilles Schaeffer, A combinatorial interpretation of super-Catalan numbers of order two, Manuscript, 2003, 4pp.

[5] Nicholas Pippenger and Kristin Schleich, Topological Characteristics of Random Triangulated Surfaces, *Random Structures & Algorithms*, to appear, arXiv:gr-qc/0306049, 2003, 58 pp.

[6] David Callan, Two bijections for Dyck path parameters, math.CO/0406381, 2004, 4pp.

2000 *Mathematics Subject Classification*: Primary 05A19; Secondary 05A15.
Keywords: super-Catalan, aligned cubic tree.

(Concerned with sequences A000108, A001700, and A007054.)

Received February 1 2005; revised version received March 2 2005. Published in *Journal of Integer Sequences*, March 2 2005.

Return to Journal of Integer Sequences home page.