A BIJECTION BETWEEN DIRECTED COLUMN-CONVEX POLYOMINOES AND ORDERED TREES OF HEIGHT AT MOST THREE

EMERIC DEUTSCH AND HELMUT PRODINGER

Abstract. A bijection is given between the set of directed column-convex polyominoes of area n and the set of ordered trees of height at most three and having n edges. Additional bijections with less well known combinatorial objects are sketched.

1. Introduction. By a *directed polyomino* we mean a polyomino that can be built by starting with a single cell and adding new cells on the right or on the top of an existing cell. By a *directed column-convex polyomino* (DCCP) we mean a directed polyomino in which every column is formed by contiguous cells. The *area* of a polyomino is defined to be the number of its cells. It is known [1] that the number of directed column-convex polyomino of area n is the Fibonacci number F_{2n-1} with $F_0 = 0, F_1 = 1$ (it is sequence A001519 in Sloane's On-Line Encyclopedia of Integer Sequences). Introducing also the empty polyomino of area zero, with count 1, it follows that the generating function for the number of directed column-convex polyominoes according to area is $\frac{1-2z}{1-3z+z^2}$. We give a simple derivation of this in the Appendix.

By the *size* of a tree we mean the number of its edges and by *height* we mean the number of edges on a maximal path starting at the root (some authors replace "edges" by "nodes" in these definitions, resulting in a difference of one unit). The *level* of a node in a tree is the number of edges on the unique path from the root to that node. It is easy to show, using for example the so-called symbolic method [5, 7], that the generating function for ordered trees of height at most three, with respect to number of edges, is again $\frac{1-2z}{1-3z+z^2}$ (see also [3, 6]). For the sake of completeness, we derive this in the Appendix.

Consequently, the number of directed column-convex polyominoes of area n is equal to the number of ordered trees of height at most three and having n edges. We present a bijection between these two sets of combinatorial objects.

2. The Bijection. For convenience, we define a *bouquet* of size k ($k = 1, 2, \ldots$) to be an ordered tree whose root has a unique child from which $k-1$ edges are emanating. In other words, a bouquet is a tree of height at most 2 and root of degree 1.

1991 *Mathematics Subject Classification.* 05A15.

Key words and phrases. Directed column-convex polyominoes; Ordered trees.

Let π be a directed column-convex polyomino of area n ($n \geq 1$). Suppose that π has width w and assume that the ith column extends from level a_i to level b_i ($i = 1, 2, \ldots, w$; $a_1 = 0$). To the directed column-convex polyomino π we associate an ordered tree τ defined in the following manner:

(i) the root of τ has degree b_w (i. e. the altitude of the top of the last column of π);

(ii) labeling the children of the root from left to right by $0, 1, 2, \ldots, b_w - 1$, on the nodes labeled $a_2, a_3, \ldots, a_w$ we hang bouquets of sizes $b_1 - a_2, b_2 - a_3, \ldots, b_{w-1} - a_w$, respectively. Note that the numbers $a_2, a_3, \ldots, a_w$ need not be distinct and, consequently, several bouquets may be dangling from the same node.

We show that the tree τ has n edges. By definition, the number of edges at level 1 is b_w and the number of nodes at level 2 and 3 is the number of edges in all the bouquets, i. e. $\sum_{i=1}^{w-1}(b_i - a_{i+1})$. Thus, the total number of edges of τ is equal to $\sum_{i=1}^{w}(b_i - a_i)$, i. e. the area n of π.

It is easy to see what is the inverse mapping. We shall indicate it on the considered example.

Example. Consider the directed column-convex polyomino π of Fig. 1. We have $w = 8$ and the values of $a_i, b_i, b_{i-1} - a_i$ are given in Table 1.

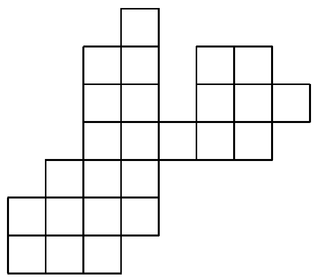

FIGURE 1. A directed column-convex polyomino of area 25

i	1	2	3	4	5	6	7	8
a_i	0	0	0	1	3	3	3	4
b_i	2	3	6	7	4	6	6	5
$b_{i-1} - a_i$		2	3	5	4	1	3	2

TABLE 1

The tree τ corresponding to π has root of degree $b_8 = 5$, the children of the root are labeled from left to right by $0, 1, 2, 3, 4$ and on the nodes with labels $0, 0, 1, 3, 3, 3, 4$ we hang bouquets of sizes, $2, 3, 5, 4, 1, 3, 2$, respectively (see Fig. 2). Conversely, assume that the tree in Fig. 2 is given. Since bouqets of sizes $2, 3, 5, 4, 1, 3, 2$ are hanging from the nodes $0, 0, 1, 3, 3, 3, 4$ of level 1, it follows that the bottoms of the columns are at levels $0, 0, 0, 1, 3, 3, 3, 4$, while the tops are at levels $2+0, 3+0, 5+1, 4+3, 1+3, 3+3, 2+4, 5$ (the last being the degree of the root).

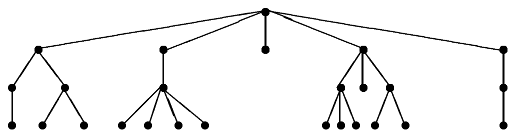

FIGURE 2. An ordered tree with 25 edges and height 3

3. Bijections with other combinatorial objects. In this section we sketch bijections between DCCPs, ordered trees of height at most three, and three other combinatorial objects. Some of these bijections reproduce those in [6] but this time we word them differently (for example, without *words*).

By a *nondecreasing Dyck path* we mean a Dyck path for which the sequence of the altitudes of the valleys is nondecreasing. They have been introduced and investigated in [2]. There is a simple bijection between DCCPs of area n and nondecreasing Dyck paths of length $2n$. We explain it on an example. The columns of the DCCP of Fig. 1, traversed from left to right, start at the altitudes $0, 0, 0, 1, 3, 3, 3, 4$ and end at the altitudes $2, 3, 6, 7, 4, 6, 6, 5$. We associate to it a nondecreasing Dyck path for which the altitudes of the valleys, including the origin, are $0, 0, 0, 1, 3, 3, 3, 4$ and the altitudes of the peaks are $2, 3, 6, 7, 4, 6, 6, 5$. We obtain the nondecreasing Dyck path of Fig. 3. The fact that the polyomino is directed ensures that the first of these sequences is nondecreasing. The inverse mapping is obvious.

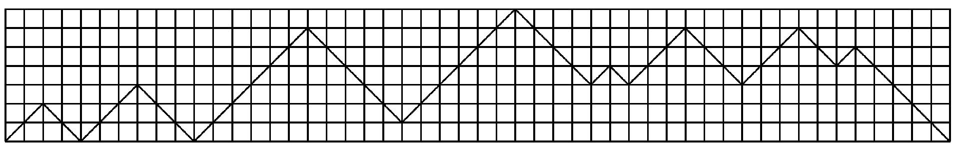

FIGURE 3. A nondecreasing Dyck path of length 50

There is a well-known bijection between Dyck paths and ordered trees (traverse the tree in preorder, the root being at the top; to each edge passed on the way down there correspond a NE step and to each edge passed on the way up there corresponds a SE step). The nondecreasing restriction on the Dyck path implies that the ordered tree can have branch nodes only on its righmost path. We call such trees *Elena trees*, or simply *Elenas* (see [6]). For example, to the nondecreasing Dyck path of Fig. 3 there corresponds the Elena of Fig. 4. For the purpose of further identification, we have marked the edges along the rightmost path by a, b, c, d, e.

We can view the Elena of Fig. 4 as a concatenation of the trees in Fig. 5. We consider these trees rooted at the top. Concatenating them by identifying their roots,

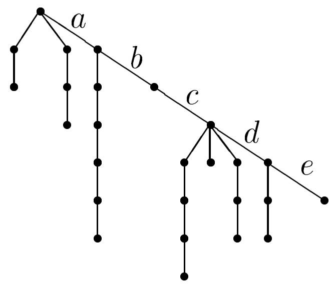

FIGURE 4. An Elena with 25 edges

we obtain the tree of Fig. 6. This is a tree with no branch nodes at height greater than one.

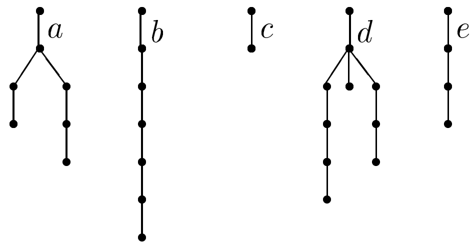

FIGURE 5. The constituents of the Elena of Fig. 4

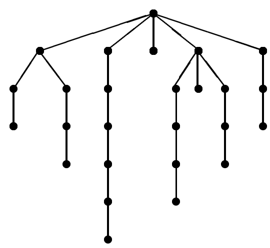

FIGURE 6. The constituents glued together at the root

Finally, if in the tree of Fig. 6 we replace all the paths that start at level 2 by a bouquet of the same size, then we obtain the tree of Fig. 2. This is a tree of height at most three.

4. Bijection correspondences.

In this section we give correspondences under the mentioned bijections between various statistics on DCCPs, nondecreasing Dyck paths, Elena trees, and ordered trees of height at most three. They follow more or less immediately from the definitions of the bijections. First we introduce some additional terminology.

In a Dyck path by a *return* we mean a SE step hitting the x-axis; by an *ascent* we mean a maximal string of NE steps. A *pyramid* in a Dyck path is a succession of NE steps followed immediately by the same number of SE steps. An *exterior pair* in a Dyck path is a pair consisting of a NE step and its matching SE step (i.e. connectable to the NE step by a horizontal line lying strictly under the path) which do not belong to any pyramid. The statistic "number of exterior pairs" and the closely related statistic "pyramid weight" have been investigated in [4]. It is easy to see that in a nondecreasing Dyck path the number of exterior pairs is equal to the altitude of the last valley.

In an ordered tree by a *branch node* we mean a node of outdegree at least 2. By a *branch* we mean a minimal succession of edges joining either the root with a branch node, or two branch nodes, or a branch node with a leaf.

In an ordered tree of height at most three we label the children of the root from left to right by 0,1,2,... (these labels will be used below).

In order to have a simpler formulation of the bijection correspondences, we assume that the considered DCCP does not consist of a single column. To a single-column DCCP of height n there correspond a nondecreasing Dyck path that is a pyramid of height n, an Elena tree that is a path of length n, and a tree of height 1 with n edges.

DCCPs	Nondecreasing Dyck paths	Elena trees	Trees of height ≤ 3
area	semilength	size	size
number of columns	number of peaks	number of leaves	1+ number of nodes at level 2
number of cells in bottom row	number of returns	root degree	total degree of the leftmost child of the root
number of cells in 1st column	altitude of 1st peak	level of leftmost leaf	label of the leftmost child of the root on which a bouquet hangs + size of this bouquet
number of cells in last column	length of rightmost ascent	length of last branch	1+ number of final leaves at level 1
level of the bottom of the last column	altitude of last valley (=number of exterior pairs)	level of last branchpoint	label of the last nonleaf child of the root
level of the top of the last column	altitude of the last peak	level of rightmost leaf	root degree

TABLE 2. How 7 parameters translate under the bijections discussed in the paper. "Last" in the columns related to trees is with respect to preorder traversal. The numerical values for our running example are 25, 8, 3, 2, 1, 4, 5, respectively.

5. Appendix. Making use of the symbolic method [7, 5], we derive (a) the generating function of the DCCPs and (b) the generating functions of ordered trees of height at most three.

(a) Every DCCP is either (i) the empty DCCP or (ii) its bottom row consists of a single cell (in which case, above it, left justified, we have a possibly empty DCCP), or (iii) to the right of its first column we have a nonempty DCCP. (see Fig. 7). Denoting the generation function by $G(z)$, this leads at once to the equation

$$G = 1 + zG + \frac{z}{1-z}(G-1), \tag{1}$$

the factor $\frac{z}{1-z}$ being the generating function of the sequence of possible first columns in the case (iii). The solution of (1) is $\frac{1-2z}{1-3z+z^2}$.

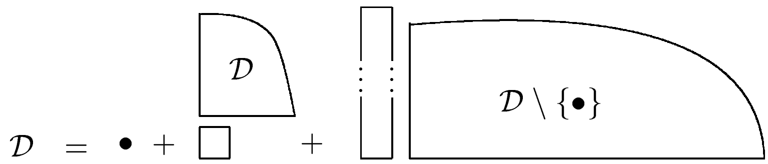

FIGURE 7. The symbolic equation for $\mathcal{D}$ =DCCP

(b) The trees of height at most 1 have generating function $\frac{1}{1-z}$. Consequently, trees of height at most 2 and having root of degree 1 have generating function $\frac{z}{1-z}$. Taking finite sequences of the latter and joining them at their roots we obtain all the trees of height at most 2 and the generating function is $\frac{1}{1-\frac{z}{1-z}} = \frac{1-z}{1-2z}$. Consequently, trees of height at most 3 and having root of degree 1 have generating function $\frac{z(1-z)}{1-2z}$ and, taking again finite sequences of the latter and joining them at their roots, we obtain all the trees of height at most three and the generating function is

$$\frac{1}{1-\dfrac{z(1-z)}{1-2z}} = \frac{1-2z}{1-3z+z^2}.$$

REFERENCES

[1] E. Barcucci, R. Pinzani, and R. Sprugnoli, Directed column-convex polyominoes by recurrence relations, in: M.-C. Gaudel and J.-P. Jouannaud (eds.), TAPSOFT '93: Theory and Practice of Software Development, Proceedings of the 4th International Joint Conference CAAP/FASE, Orsay, France, April 1993, Lecture Notes in Computer Science, No. 668, Springer, Berlin (1993), pp. 282-298.

[2] E. Barcucci, A. Del Lungo, S. Fezzi, and R. Pinzani, Nondecreasing Dyck paths and q-Fibonacci numbers, Discrete Math., 170, 211-217, 1997.

[3] N. G. de Bruijn, D. E. Knuth, and S. O. Rice, The average height of planted trees, in: Graph Theory and Computing (edited by R. C. Read), Academic Press, New York (1972), pp. 15-22. Reprinted in: Selected papers on analysis of algorithms. CSLI Publications, Stanford, CA, 2000. xvi+621 pp.

[4] A. Denise and R. Simion, Two combinatorial statistics on Dyck paths, Discrete Math., 137, 155-176, 1995.

[5] P. Flajolet and R. Sedgewick, Analytic Combinatorics. Book in preparation, 1998. (Individual chapters are available as INRIA Research Reports 1888, 2026, 2376, 2956, 3162.)

[6] H. Prodinger, Words, Dyck paths, trees, and bijections. In: Words, semigroups, and transductions, 369–379, World Sci. Publishing, River Edge, NJ, 2001.

[7] R. Sedgewick and P. Flajolet, An Introduction to the Analysis of Algorithms, Addison-Wesley, Reading, 1996.

[8] N. J. A. Sloane and S. Plouffe, The Encyclopedia of Integer Sequences, Academic Press, San Diego, 1995.

[9] R. P. Stanley, Enumerative Combinatorics, Vol. 2, Cambridge University Press, Cambridge, 1999.

Department of Mathematics, Polytechnic University, Brooklyn, NY 11201
E-mail address: `deutsch@duke.poly.edu`

The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of the Witwatersrand, P. O. Wits, 2050 Johannesburg, South Africa
E-mail address: `helmut@maths.wits.ac.za`