Generating Trees and Proper Riordan Arrays

Donatella Merlini, M. Cecilia Verri
Dipartimento di Sistemi e Informatica
via Lombroso 6/17, 50134, Firenze, Italia
merlini,verri@dsi.unifi.it

Abstract

We use an algebraic approach to study the connection between *generating trees* and *proper Riordan Arrays* deriving a theorem that, under suitable conditions, associates a Riordan Array to a generating tree and vice versa. Thus, we can use results from the theory of Riordan Arrays to study properties of generating trees. In particular, we can find, in a general and easy way, the generating functions counting the distribution of trees' nodes at each level. The connection between Riordan Arrays and transfer matrices is also shown.

1 Introduction

The concept of *generating trees* has been introduced in the literature by Chung, Graham, Hoggat and Kleiman in [2] to examine the reduced Baxter permutations. This technique has been successfully applied by West [11, 12] to other classes of permutations and more recently to some other combinatorial classes such as plane trees and lattice paths (see, Barcucci, Del Lungo, Pergola and Pinzani [1]). In all these cases, a generating tree is associated to a certain combinatorial class, according to some enumerative parameter, in such a way that the number of nodes appearing on level n of the tree gives the number of n-sized objects in the class.

If a problem has been defined by means of a generating tree, some device has to be used to obtain counting information on the objects of the associated combinatorial class. In this paper, we introduce the concept of *matrix associated to a generating tree* (AGT matrix, for short): this is an infinite matrix $\{d_{n,k}\}_{n,k\in\mathbf{N}}$ where $d_{n,k}$ is the number of nodes at level n with label $k+c$, c being the label of the root. Our main result is Theorem 3.9 which states the conditions under which an AGT matrix is a *proper Riordan Array*, and vice versa. The concept of a Riordan Array provides a remarkable characterization of many lower triangular arrays that arise in combinatorics and algorithm analysis. The theory has been introduced in the literature in 1991 by Shapiro, Getu, Woan and Woodson [9] and then examined closely from a theoretical and practical point of view by Merlini, Rogers, Sprugnoli and Verri [4, 5, 10]. This study has pointed out that Riordan Arrays are a powerful tool in the study of many counting problems and Theorem 3.9 is a further interesting application of this approach. Moreover, Theorem 3.10, illustrates the relation between proper Riordan Arrays and the transfer matrices.

The structure of the paper is as follows. In Section 2 we define the concepts of generating trees and AGT matrices and show some applications in which generating trees are used. In Section 3 we summarize the main results concerning Riordan Arrays and prove Theorems 3.9 and 3.10. For the sake of completeness, we then propose some generating trees which do not correspond to Riordan Arrays and, vice versa, some Riordan Arrays not directly related to generating trees. In Section 4, we give some applications of Theorem 3.9. First, we show how some well-known generating trees can be studied from the point of view of Riordan Arrays and thus how counting results are obtained. Finally, we take into considerations some lattice path problems (which have been studied as Riordan Arrays in a direct way by Merlini, Sprugnoli and Verri in [5]) and prove that they can be defined as generating trees.

2 Generating trees

The concept of generating trees has been used from various points of view and, as we noted in the introduction, has been introduced in the literature by Chung, Graham, Hoggat and Kleiman [2] to examine the reduced Baxter permutations. This technique has been successively applied to other classes of permutations and the main references on the subject are due to West [11, 12]. We begin by giving the following:

Definition 2.1 *A generating tree is a rooted labelled tree with the property that if v_1 and v_2 are any two nodes with the same label then, for each label l, v_1 and v_2 have exactly the same number of children with label l. To specify a generating tree it therefore suffices to specify:*

1) the label of the root;

2) a set of rules explaining how to derive from the label of a parent the labels of all of its children.

For example, Figure 2.1 illustrates the upper part of the generating tree which corresponds to the following specification:

$$\begin{cases} root: & (2) \\ rule: & (k) \quad \rightarrow (2)\cdots(k)(k+1) \end{cases} \tag{2.1}$$

We associate to a generating tree an infinite matrix $\{d_{n,k}\}_{n,k\in\mathbf{N}}$ defined as follows:

Definition 2.2 *An infinite matrix $\{d_{n,k}\}_{n,k\in\mathbf{N}}$ is said to be "associated" to a generating tree with root (c) (AGT matrix for short) if $d_{n,k}$ is the number of nodes at level n with label $k+c$. By convention, the level of the root is* 0.

Referring to rule (2.1) and to Figure 2.1 we have the following partial associated matrix:

n/k	0	1	2	3	4
0	1				
1	1	1			
2	2	2	1		
3	5	5	3	1	
4	14	14	9	4	1

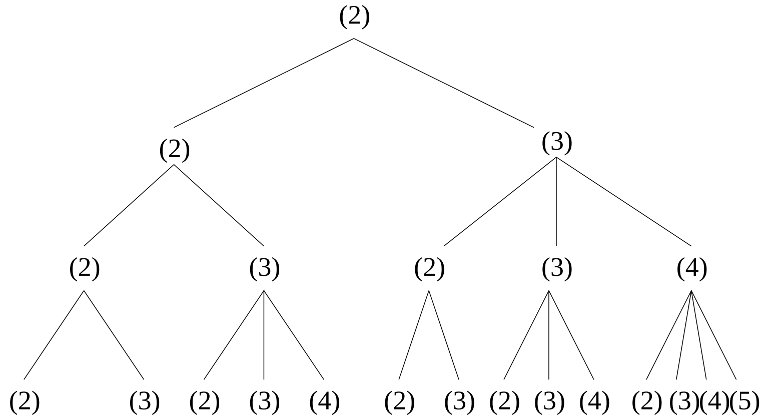


Figure 2.1: The partial generating tree for the specification (2.1).

where we recognize the Catalan triangle.

We observe that an AGT matrix is always infinite in the direction of n but can be finite in the direction of k because the node's labels of the generating tree can be all less than r for some $r \in \mathbf{N}$. For example, this is the case of the specification:

$$\begin{cases} \mathit{root}: & (1) \\ \mathit{rule}: & (1) \rightarrow (2);\ (2) \rightarrow (1)(2) \end{cases} \tag{2.2}$$

for which we have $d_{n,k} = 0\ \forall k \geq 2$. In these cases, the generating tree specification can be seen as the homomorphism $h : \Sigma^* \rightarrow \Sigma^*$ associated to a D0L system $G = (\Sigma, h, \alpha)$ (see Rozenberg and Salomaa [7] for reference) where Σ is a finite alphabet and $\alpha \in \Sigma^*$ is the "axiom" ($\Sigma = \{1, 2\}$ $h(1) = 2$, $h(2) = 12$ and $\alpha = 1$, in the previous example); system G generates the language L_G obtained by iteratively applying the homomorphism h to α : $L_G = \{\alpha, h(\alpha), h^2(\alpha), \ldots\}$.

Two interesting quantities related to generating trees are the total number s_n of nodes at level n and the average label v_n of a node at the same level n. In fact, these quantities have an obvious combinatorial meaning. By going on to the AGT matrix D associated to the generating tree, we immediately see that s_n is given by the row sums of D and v_n is the ratio between the weighted row sums W_n of D and s_n, plus c:

$$s_n = \sum_{k=0}^{n} d_{n,k}, \qquad v_n = \frac{\sum_{k=0}^{n}(k+c)d_{n,k}}{s_n} = \frac{\sum_{k=0}^{n} k d_{n,k}}{s_n} + c = \frac{W_n}{s_n} + c.$$

As we will see, s_n and W_n can be easily computed if the AGT matrix is a Riordan Array. In fact, the Riordan Array theory allows us to compute many combinatorial quantities related to generating trees.

West [12] discusses how generating trees can be used to enumerate some classes of permutations with forbidden subsequences and argues for a broader use of generating trees in combinatorial enumeration. We now summarize West's technique and refer back to his paper [12] for a complete treatment. If S_n denotes the set of permutations on $\{1, 2, \ldots, n\}$ and $\tau \in S_k$, $k \leq n$, a permutation $\pi \in S_n$ is τ-avoiding iff there is no $1 \leq i_{\tau(1)} < i_{\tau(2)} < \ldots i_{\tau(k)} \leq n$ such that $\pi(i_1) < \pi(i_2) \ldots < \pi(i_k)$; $S_n(\tau)$ denotes the set of all τ-avoiding permutations of length n. Given τ, he associates a generating tree to the set of τ avoiding permutations by defining a rooted tree in which the nodes on level n are precisely the elements of $S_n(\tau)$. The tree is constructed from the root down by inserting n in every position where it does not create a τ-subsequence. The node's labels correspond to the number of children each node has. For example, if $\tau = 123$, a permutation $\pi \in S_{n-1}(123)$, $\pi = p_1, \ldots, p_{n-1}$, has as potential children the permutations $\pi^j = p_1, \ldots, p_{j-1}, n, p_j, \ldots, p_{n-1}$, as j ranges from 1 to n. In particular, π has exactly k children if its longest initial decreasing subsequence (i.d.s.) has length $k-1$. And looking at its children, π^1 has an i.d.s. of length k while π^j, $2 \leq j \leq k$, has an i.d.s. of length $j-1$. This means, in summary, that the node's labels follow rule (2.1). To count permutations he rewrites the generating tree specification in terms of a transition matrix T such that $T_{i,j}$ is equal to α if $(j) \to (i)^\alpha$ and is 0 otherwise. For example, for the specification (2.2), he finds:

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

If the transition matrix is T the vector giving the n-th level numbers is $T^n e_c^\top$, where e_c is the unit vector having a 1 in the c-th position and 0s elsewhere. In the previous example he thus obtains:

$$T^n \binom{1}{0} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \binom{1}{0} = \binom{F_{n-2}}{F_{n-1}}$$

where F_n is the n^{th} Fibonacci number with initial conditions $F_0 = F_1 = 1$. Some other complex relations between permutations and generating trees are also illustrated.

More recently, Barcucci, Del Lungo, Pergola and Pinzani in [1] formalize a methodology for the enumeration of various combinatorial classes making use of the concept of generating trees. Their basic idea is the following: given a class S of combinatorial objects, let S_n be the set of n-sized objects of S, according to some parameter. For example, if S is the class of plane trees we can take S_n as the set of plane trees with n internal nodes, each connected to exactly one external node. Moreover, let $\theta : S_n \to 2^{S_{n+1}}$ be an operator with the following properties:

1) for each $Y \in S_{n+1}$ there exists $X \in S_n$ such that $Y \in \theta(X)$;

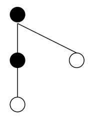
2) if $X_1, X_2 \in S_n$ and $X_1 \neq X_2$ then $\theta(X_1) \cap \theta(X_2) = \emptyset$.

Then, the family of sets $\{\theta(X) : X \in S_n\}$ is a partition of S_{n+1}. This result allows us to construct each object $Y \in S_{n+1}$ from an object $X \in S_n$ in a unique way and thus θ yields a recursive description of the class S. For example, for the class of plane trees cited above, the operator θ which replaces, with an internal node, each external node that follows the last internal node in the preorder traversal, satisfies properties 1) and 2).

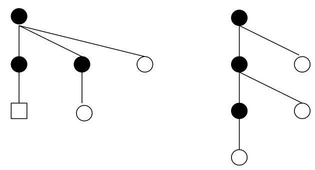
An *active site* is a place where an object can be expanded. Given an operator θ satisfying conditions 1) and 2) and $X \in S$, let $\mathcal{F}(X)$ denotes the set of X's active sites for θ, (in the previous example this is the set of external nodes that follow the last internal node in the preorder traversal). If $|\mathcal{F}(X)| = k$, let X_i, $1 \leq i \leq k$, denotes the object obtained from X by expanding the i-th element in $\mathcal{F}(X)$. Finally, a generating tree can be associated to the class S by means of the following specification:

$$\begin{cases} root: & (c) \\ rule: & (k) \quad \to (q_1) \cdots (q_k) \end{cases} \tag{2.3}$$

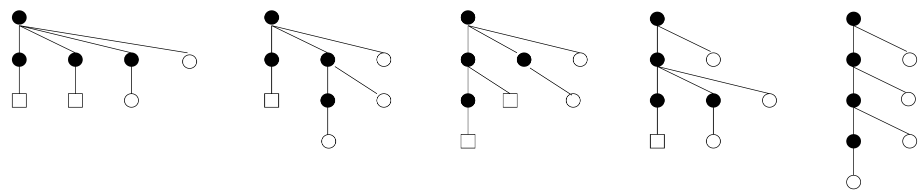
Here c is the number of active sites for the object that gives rise to the recursive process and if $X \in S$ has $|\mathcal{F}(X)| = k$ then we have $q_i = |\mathcal{F}(X_i)|, 1 \leq i \leq k$. For example, in the plane tree case, it can be proved that the specificaton rule is given by (2.1). Since the application of the θ operator to an object of size n and with i active sites gives i objects of size $n+1$ then the AGT matrix $\{d_{n,k}\}_{n,k\in\mathbf{N}}$ associated to the generating tree (2.1) has the following combinatorial interpretation: $d_{n,k}$ is the number of objects of size $n+n_0$ having $k+c$ active sites, n_0 being the size of the object that gives rise to the process. Moreover, s_n is the total number of objects of size $n+n_0$ and v_n is the average value of active sites that an object of size $n+n_0$ has. In the plane tree case described above, let us indicate an internal node with ● , an external node which is an active site with ○ and an external node which is not an active site with □ ; if we start from the plane tree



which corresponds to $d_{0,0} = 1$, then we obtain the two plane trees



which correspond to $d_{1,0} = 1$ and $d_{1,1} = 1$. Successively we find



which correspond to $d_{2,0} = 2$, $d_{2,1} = 2$, and $d_{2,2} = 1$, and so on.

In Barcucci, Del Lungo, Pergola and Pinzani [1] various generating trees are taken into consideration and for each of them the relation with some well-known combinatorial structures is shown.

3 The relation with Riordan Arrays

The concept of a *Riordan array* has been introduced in 1991 by Shapiro, Getu, Woan and Woodson [9] (they chose this name in honour of John Riordan), with the aim of generalizing the concept of *Renewal Array* defined by Rogers [6] in 1978. Their basic idea was to define a class of infinite lower triangular arrays with properties analogous to those of the Pascal triangle whose elements, as is well-known, are the binomial coefficients $\binom{n}{k}$.

This concept has also been studied by Sprugnoli [10], who pointed out the relevance of these matrices from a theoretical and practical point of view. Successively, some other aspects of the theory have been studied and we cite here Merlini, Rogers, Sprugnoli and Verri [4, 5]).

A Riordan Array is an infinite lower triangular array $\{d_{n,k}\}_{n,k\in\mathbf{N}}$, defined by a pair of formal power series $D = (d(t), h(t))$, such that the generic element $d_{n,k}$ is the n-th coefficient in the series $d(t)(th(t))^k$, i.e.:

$$d_{n,k} = [t^n]d(t)(th(t))^k, \qquad n, k \geq 0.$$

From this definition we have $d_{n,k} = 0$ for $k > n$:

In the sequel we always assume that $d(0) \neq 0$; if we also have $h(0) \neq 0$ then the Riordan array is said to be *proper*; in the proper-case the diagonal elements $d_{n,n}$ are different from zero for all $n \in \mathbf{N}$. The most simple example is the Pascal triangle for which we have

$$\binom{n}{k} = [t^n]\frac{1}{1-t}\left(\frac{t}{1-t}\right)^k,$$

where we recognize the proper Riordan Array $d(t) = h(t) = 1/(1-t)$.

We shall now describe the most important results concerning Riordan Arrays and refer back to Merlini, Rogers, Sprugnoli and Verri [4] for the complete theory and for the proofs.

We begin with the following:

Theorem 3.1 *Let $D = (d(t), h(t))$ be a Riordan Array and $f(t)$ the generating function for the sequence $\{f_k\}_{k\in\mathbf{N}}$. Then:*

$$\sum_{k=0}^{n} d_{n,k} f_k = [t^n]d(t)f(th(t)).$$

As noted in Section 2, two important quantities are the row sums and the weighted row sums of an array, If we apply Theorem 3.1 with $f_k = 1$ or $f_k = k$, $\forall k \in \mathbf{N}$, we have $f(t) = 1/(1-t)$ and $f(t) = t/(1-t)^2$ respectively, hence:

$$s_n = \sum_{k=0}^{n} d_{n,k} = [t^n]\frac{d(t)}{1-th(t)}, \qquad W_n = \sum_{k=0}^{n} kd_{n,k} = [t^n]\frac{td(t)h(t)}{(1-th(t))^2}.$$

Once we have these generating functions we can extract the exact coefficient or find an asymptotic approximation in order to have the sum's value.

Proper Riordan Arrays are characterized by the following fundamental property found by Rogers [6] in 1978 and then examined closely by Sprugnoli [10].

Theorem 3.2 *A matrix $\{d_{n,k}\}_{n,k\in\mathbf{N}}$ is a proper Riordan Array iff there exists a sequence $A = \{a_i\}_{i\in\mathbf{N}}$ with $a_0 \neq 0$ s.t. every element $d_{n+1,k+1}$ can be expressed as a linear combination, with coefficients in A, of the elements in the preceding row, starting from the preceding column:*

$$d_{n+1,k+1} = a_0 d_{n,k} + a_1 d_{n,k+1} + a_2 d_{n,k+2} + \cdots$$

The previous sum is finite since $d_{n,k} = 0$ for $k > n$. The sequence A is called the *A-sequence* of the Riordan Array and is characteristic of the matrix since it determines the function $h(t)$ and vice versa. In fact we have the following:

Corollary 3.3 *Let $D = (d(t), h(t))$ be a proper Riordan Array, and let $A = \{a_j\}_{j\in\mathbf{N}}$ be its A-sequence. Then, if $A(t)$ is the generating function of the sequence A, we have:*

$$h(t) = A(th(t)).$$

For example, for the Pascal triangle we have:

$$A(y) = \left[\frac{1}{1-t}\middle| y = \frac{t}{1-t}\right] = \left[\frac{1}{1-t}\middle| t = \frac{y}{1+y}\right] = 1 + y,$$

and thus the A-sequence for this triangle is $\{1, 1, 0, 0, \ldots\}$. The relation of Theorem 3.2 reduces to the well-known recurrence relation for binomial coefficients:

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

Two other useful results are given by the following theorem and corollary:

Theorem 3.4 *Let $D = (d(t), h(t))$ be a proper Riordan Array, and let $A = \{a_j\}_{j\in\mathbf{N}}$ be its A-sequence. Then, if $d(t) = 1$ we have:*

$$d_{n,k} = \frac{k}{n}[t^{n-k}]A(t)^n,$$

and if $d(t) = h(t)$ we have:

$$d_{n,k} = \frac{k+1}{n+1}[t^{n-k}]A(t)^{n+1}.$$

Corollary 3.5 *Let $h(t)$ be a formal power series with $h(0) \neq 0$ and $f_n = [t^n]h(t)^p$ for some $p \in \mathbf{N}$. Then we have:*

$$f_n = \frac{p}{p+n}[t^n]A(t)^{n+p}$$

where $h(t) = A(th(t))$.

More recently some new aspects of the Riordan Array theory have been studied (see Merlini, Rogers, Sprugnoli and Verri [4]). The A-sequence doesn't characterize completely $(d(t), h(t))$ because $d(t)$ is independent of $A(t)$. But we have the following:

Theorem 3.6 *Let $\{d_{n,k}\}_{n,k\in\mathbf{N}}$ be an infinite lower triangular array with $d_{n,n} \neq 0, \forall n \in \mathbf{N}$ (in particular, let it be a proper Riordan Array); then there exists a unique sequenze $Z = \{z_0, z_1, z_2, \ldots\}$ such that every element in column 0 can be expressed as a linear combination of all the elements of the preceding row:*

$$d_{n+1,0} = z_0 d_{n,0} + z_1 d_{n,1} + z_2 d_{n,2} + \cdots$$

The Z-sequence characterizes column 0 while the A-sequence characterizes all the other columns. We can conclude that the triple $(d_0, Z(t), A(t))$, with $d_0 = d(0)$, characterizes every proper Riordan Array:

Theorem 3.7 *Let $(d(t), h(t))$ be a proper Riordan Array and let $Z(t)$ be the generating function for the Z-sequence of the matrix. Then we have:*

$$d(t) = \frac{d_0}{1 - tZ(th(t))}.$$

The previous relation can be inverted and this allows us to find a formula for the Z-sequence:

$$Z(y) = \left[\frac{d(t) - d_0}{td(t)} \middle| t = yh(t)^{-1} \right].$$

For example, in the Pascal triangle, we have:

$$Z(y) = \left[\left(\frac{1}{1-t} - 1 \right) \frac{1-t}{t} \middle| y = \frac{t}{1-t} \right] = \left[1 \middle| t = \frac{y}{1+y} \right] = 1.$$

In fact, the Pascal triangle belongs to the following general case:

Theorem 3.8 *Let $d_0 = h_0 \neq 0$. Then $d(t) = h(t)$ iff $A(y) = d_0 + yZ(y)$.*

Now we have all the theoretical tools necessary to study the connection between proper Riordan Arrays and generating trees. More precisely, the following theorem gives the conditions under which an AGT matrix associated to a generating tree is a Riordan Array and vice versa. In its statement, we introduce a special *product notation* which plays an important role in the proof and in the application of the same theorem.

Theorem 3.9 *Let $c \in \mathbf{N}$, $a_j \in \mathbf{N}$, $a_0 \neq 0$, $b_k \in \mathbf{Z}$, and $b_{c+j} + a_{j+1} \geq 0$, $\forall j \geq 0$ and $k \geq c$, and let*

$$\begin{cases} root: & (c) \\ rule: & (k) \quad \rightarrow (c)^{b_k} \prod_{j=0}^{k+1-c} (k+1-j)^{a_j} \end{cases} \tag{3.4}$$

be a generating tree specification. Then, the AGT matrix associated to (3.4) is a proper Riordan Array D defined by the triple (d_0, A, Z), such that

$$d_0 = 1, \quad A = (a_0, a_1, a_2, \ldots), \quad Z = (b_c + a_1, b_{c+1} + a_2, b_{c+2} + a_3, \ldots).$$

Vice versa, if D is a proper Riordan Array defined by the triple (d_0, A, Z) with $d_0 = 1$ and $a_j, z_j \in \mathbf{N}$, $\forall j \geq 0$, then D is the AGT matrix associated to the generating tree specification (3.4) with $b_{c+j} = z_j - a_{j+1}$, $\forall j \geq 0$.

Proof: Let us consider the AGT matrix $D = \{d_{n,k}\}_{n,k \in \mathbf{N}}$ associated to (3.4). Then, $d_{n,k}$ counts the number of nodes at level n with label $k+c$. We have obviously $d_{0,0} = 1$. Moreover, we observe that the maximum label's value at each level increases by one with respect to the previous level, hence $d_{n,j} = 0$ for $j > n$. Now, (3.4) tells us that a node at level $n+1$ with label $k+1+c$ can be determined, in a_j different ways, from the nodes at level n with label $h+c$, such that $h+c+1-j = k+c+1$, i.e., $h = k+j$, $j \geq 0$, hence:

$$d_{n+1,k+1} = a_0 d_{n,k} + a_1 d_{n,k+1} + \cdots$$

Finally, a node at level $n+1$ with label c can be obtained in $b_h + a_{h+1-c}$ different ways from the nodes at level n with label h, $h \geq c$. Hence we have:

$$d_{n+1,0} = (b_c + a_1)d_{n,0} + (b_{c+1} + a_2)d_{n,1} + \cdots$$

This proves the "if" part of the theorem.

Conversely, we only need to observe that the elements of an AGT matrix are all positive and this is assured for the elements of a Riordan Array having $d_0 = 1$ and $a_j, z_j \in \mathbf{N}$, $\forall j \geq 0$. ■

Another important result is the following theorem that connects the concept of proper Riordan Arrays with the one of transfer matrices used by West [12], pointing out the power of the Riordan Array approach:

Theorem 3.10 *Let $D = \{d_{n,k}\}_{n,k\in\mathbf{N}}$ be a proper Riordan Array defined by the triple (d_0, A, Z) and let T and $\bar{d}_0$ be an infinite matrix and an infinite vector defined as follow:*

$$T = \begin{pmatrix} z_0 & z_1 & z_2 & z_3 & z_4 & \ldots \\ a_0 & a_1 & a_2 & a_3 & a_4 & \ldots \\ 0 & a_0 & a_1 & a_2 & a_3 & \ldots \\ 0 & 0 & a_0 & a_1 & a_2 & \ldots \\ 0 & 0 & 0 & a_0 & a_1 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \bar{d}_0 = \begin{pmatrix} d_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

Then, the generic element $d_{n,k}$ of the Riordan Array, $n > 0$, is the k^{th} element of the vector $T^n\bar{d}_0$.

Proof: By induction on n. For $n = 1$ we have $(T\bar{d}_0)^\top = (z_0d_0, a_0d_0, 0, \ldots)$ and obviously $d_{1,0} = z_0d_0$ and $d_{1,1} = a_0d_0$. Suppose that $(T^n\bar{d}_0)^\top$ is the n^{th} row of the Riordan Array, that is, $(T^n\bar{d}_0)^\top = (d_{n,0}, d_{n,1}, d_{n,2}, \ldots)$, and let us look at $T^{n+1}\bar{d}_0$. We have:

$$T^{n+1}\bar{d}_0 = T\left(T^n\bar{d}_0\right) = \begin{pmatrix} z_0 & z_1 & z_2 & z_3 & z_4 & \ldots \\ a_0 & a_1 & a_2 & a_3 & a_4 & \ldots \\ 0 & a_0 & a_1 & a_2 & a_3 & \ldots \\ 0 & 0 & a_0 & a_1 & a_2 & \ldots \\ 0 & 0 & 0 & a_0 & a_1 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} d_{n,0} \\ d_{n,1} \\ d_{n,2} \\ \vdots \end{pmatrix} =$$

$$= \begin{pmatrix} z_0d_{n,0} + z_1d_{n,1} + z_2d_{n,2} + \cdots \\ a_0d_{n,0} + a_1d_{n,1} + a_2d_{n,2} + \cdots \\ a_0d_{n,1} + a_1d_{n,2} + a_2d_{n,3} + \cdots \\ \vdots \\ \end{pmatrix} = \begin{pmatrix} d_{n+1,0} \\ d_{n+1,1} \\ d_{n+1,2} \\ \vdots \end{pmatrix}.$$

The last passage is due to Theorems 3.2 and 3.6 and proves the theorem. ■

In the next section, we will give some important applications of Theorem 3.9. Here, it might be interesting to see, for the sake of completeness, some "negative" examples, in particular, two generating trees whose AGT matrix is not a proper Riordan Array and a proper Riordan Array which cannot be seen as associated to any generating tree.

In Theorem 3.9 the a_j's are independent from k. Nevertheless, if $a_j = a_j(k)$ for some j we can repeat the considerations done in the theorem's proof and find a recurrence relation between $d_{n+1,k+1}$ and the element in the preceding row. For example, the following specification:

$$\begin{cases} \mathit{root}: & (0) \\ \mathit{rule}: & (k) \quad \rightarrow (k)^k(k+1) \end{cases} \tag{3.5}$$

corresponds to $a_0 = 1$, $a_1 = k$ and $a_j = 0$ for $j \geq 1$, that is:

$$d_{n+1,k+1} = d_{n,k} + kd_{n,k+1}.$$

This is the well-known recurrence relation for Stirling numbers of the second kind $\left\{ {n \atop k} \right\}$ and we have the following partial associated matrix:

n/k	0	1	2	3	4
0	1				
1	0	1			
2	0	1	1		
3	0	1	3	1	
4	0	1	7	6	1

Actually, this is not a Riordan Array, but it is easy to prove that the matrix having $d_{n,k} = \frac{k!}{n!} \left\{ {n \atop k} \right\}$ is a proper Riordan Array defined by $d(t) = 1$ and $h(t) = \frac{e^t - 1}{t}$ (see Sprugnoli [10]).

Let us now consider the following specification:

$$\begin{cases} \mathit{root}: & (1) \\ \mathit{rule}: & (k) \quad \rightarrow (1)(2)\cdots(k-1)(k+2) \end{cases} \tag{3.6}$$

In this case we get the following partial associated matrix:

n/k	0	1	2	3	4	5	6	7	8
0	1								
1	0	0	1						
2	1	1	0	0	1				
3	2	1	2	2	0	0	1		
4	6	5	5	2	3	3	0	0	1

which obviously is not a Riordan Array. We observe explicitly that the lengthened shape of the matrix is due the presence of $(k+2)$ in (3.6). In general, when the generating tree specification contains a factor of type $(k+j)$ with $j > 1$, the associated matrix we obtain is not a proper Riordan Array. For these cases, it would be interesting and useful to investigate the connection between AGT matrices and the concept of *stretched Riordan*

Arrays, as defined by Corsani, Merlini and Sprugnoli in [3]. But we don't intend to study these aspects in this paper.

An example of a proper Riordan Array which is not an AGT matrix is given by the triple (d_0, A, Z), with $d_0 = 1$, $A = (1, -1, 1, 0, 0 \ldots)$, and $Z = (1, 1, 0, 0, \ldots)$, or, equivalently,

$$d(t) = \frac{1 - 3t - \sqrt{(1-t)(1+3t)}}{2t(3t-2)}, \qquad h(t) = \frac{1 + t - \sqrt{(1-t)(1+3t)}}{2t^2}.$$

Some of the $d_{n,k}$'s values are negative, as illustrated in the upper part of the triangle:

n/k	0	1	2	3	4
0	1				
1	1	1			
2	2	0	1		
3	2	3	-1	1	
4	5	-2	5	-2	1

Obviously, this matrix can't be associated to any generating tree since a negative element can't represent the number of nodes with a certain label at a certain level. Nevertheless, from an algebraic point of view, we can associate this Riordan Array to the following specification:

$$\begin{cases} root: & (1) \\ rule: & (1) \quad \to (1)(2);\ (k) \to (k-1)(k)^{-1}(k+1) \end{cases} \tag{3.7}$$

where the meaning of $(k)^{-1}$ will be immediately explained. This allows us to give a combinatorial interpretation of the array. In fact, starting from the root, we can build a succession of lists whose elements are of type k or $\overline{j}$ with $k, j \in \mathbf{N}$: a $\overline{j}$ element comes from the application of the rule $(j) \to (j-1)(j)^{-1}(j+1)$ and in particular corresponds to $(j)^{-1}$; on the other hand, the application of the rule to an element of type $\overline{j}$ gives elements $\overline{j-1}, j$, and $\overline{j+1}$. Finally, we delete every pair $k, \overline{k}$ in each list, obtaining the following succession:

$$\begin{array}{l}
\{1\} \\
\{1,2\} \\
\{1,2,1,\overline{2},3\} = \{1,1,3\} \\
\{1,2,1,2,2,\overline{3},4\} \\
\{1,2,1,\overline{2},3,1,2,1,\overline{2},3,1,\overline{2},3,\overline{2},3,\overline{4},3,\overline{4},5\} = \{1,1,1,1,1,\overline{2},\overline{2},3,3,3,3,3,\overline{4},\overline{4},5\} \\
\vdots
\end{array}$$

which tells us that the elements $d_{n,k}$ of the above Riordan Array correspond to the number of elements equal to $k+1$ in the n^{th} list, when we count the overlined elements as negative.

4 Some examples

In this section we describe some applications of Theorem 3.9 in two directions: 1) we illustrate some well-known generating tree specifications and then find the associated proper Riordan Arrays; this will allow us to obtain some very important combinatorial results, which are easily and directly derived in the Riordan Array approach; 2) starting from some well-known proper Riordan Arrays, we find the corresponding generating tree specifications.

Let us start by applying the theorem to the generating tree specification (2.1): we have $c = 2$, $b_k = 0$, and $a_k = 1$, $\forall k \geq 0$. We therefore obtain the proper Riordan Array $\{d_{n,k}\}_{n,k\in\mathbf{N}}$ defined by the triple (d_0, A, Z) with:

$$d_0 = 1, \quad A = (1, 1, 1, \ldots), \quad Z = (1, 1, 1, \ldots).$$

We have $A(t) = Z(t) = 1/(1-t)$ and from Corollary 3.3 and Theorem 3.8 we get:

$$d(t) = h(t), \qquad h(t) = \frac{1}{1 - th(t)}.$$

By solving the previous functional equation we find the generating function of the Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$:

$$d(t) = h(t) = 1 + th(t)^2 = \frac{1 - \sqrt{1-4t}}{2t}.$$

The generic element of the Riordan Array is given by Theorem 3.4:

$$d_{n,k} = \frac{k+1}{n+1}[t^{n-k}]\left(\frac{1}{1-t}\right)^{n+1} = \frac{k+1}{n+1}\binom{-n-1}{n-k}(-1)^{n-k} = \frac{k+1}{n+1}\binom{2n-k}{n-k}.$$

These values represent the number of nodes at level n having label $k+1$ in the rooted labelled tree described in Figure 2.1. The row and weighted row sums can be easily determined by means of Corollary 3.5:

$$\sum_{k=0}^{n} d_{n,k} = [t^n]\frac{h(t)}{1-th(t)} = [t^n]h(t)^2 =$$

$$= \frac{2}{n+2}[t^n](1-t)^{-n-2} = \frac{1}{n+2}\binom{2(n+1)}{n+1} = C_{n+1},$$

$$\sum_{k=0}^{n} kd_{n,k} = [t^n]\frac{th(t)^2}{(1-th(t))^2} = [t^{n-1}]h(t)^4 =$$

$$= \frac{4}{n+3}[t^{n-1}](1-t)^{-n-3} = \frac{2n}{(n+2)(n+3)}\binom{2(n+1)}{n+1} = \frac{2n}{n+3}C_{n+1}.$$

Therefore, the average label of a node at level n is $v_n = \frac{2n}{n+3} + 2$.

In Barcucci, Del Lungo, Pergola and Pinzani [1] we also find two other generating tree specifications which correspond to some well-known combinatorial classes and that lie in the hypothesis of Theorem 3.9:

$$\begin{cases} root: & (1) \\ rule: & (k) \rightarrow (1)(2)\cdots(k-1)(k+1) \end{cases} \tag{4.8}$$

$$\begin{cases} root: & (2) \\ rule: & (k) \rightarrow (3)(4)\cdots(k)(k+1)(k+1) \end{cases} \tag{4.9}$$

For the specification (4.8) we have $c = 1$, $b_k = 0$, $\forall k \in \mathbf{N}$ and $a_k = 0$ for $k = 1$, $a_k = 1$ otherwise. Hence, the associated AGT matrix is a proper Riordan Array defined by the triple:

$$d_0 = 1, \quad A = (1, 0, 1, 1, \ldots), \quad Z = (0, 1, 1, 1, \ldots);$$

we have again $d(t) = h(t)$ and from $A(t) = (1 - t - t^2)/(1 - t)$ it follows that:

$$d(t) = h(t) = \frac{1 + t - \sqrt{1 - 2t - 3t^2}}{2t(1 + t)} = 1 + t^2 + t^3 + 3t^4 + 6t^5 + 15t^6 + 36t^7 + O(t^8)$$

which are related to the Motzkin numbers $M_n = [t^n]\frac{1-t-\sqrt{1-2t-3t^2}}{2t}$. By developing around the dominating singularity $t = 1/3$ and taking the first order approximation, we have:

$$\sum_{k=0}^{n} d_{n,k} = [t^n]\frac{\left(1 + t - \sqrt{1 - 2\,t - 3\,t^2}\right)^2}{4t^2\,(1 + t)} \approx -[t^{n+2}]\frac{\sqrt{3}}{3}\sqrt{1 - 3t} =$$

$$= \frac{\sqrt{3}}{3(2n + 3)}\left(\frac{3}{4}\right)^{n+2}\binom{2(n + 2)}{n + 2}$$

$$\sum_{k=0}^{n} k d_{n,k} = [t^n]\frac{\left(1 - t - \sqrt{1 - 2\,t - 3\,t^2}\right)\left(1 + t - \sqrt{1 - 2\,t - 3\,t^2}\right)^2}{8t^3\,(1 + t)} \approx$$

$$\approx -[t^{n+3}]\frac{2\sqrt{3}}{9}\sqrt{1 - 3t} = \frac{2\sqrt{3}}{9(2n + 5)}\left(\frac{3}{4}\right)^{n+3}\binom{2(n + 3)}{n + 3}.$$

For the specification (4.9) we have $c = 2$, $b_k = -1$, $\forall k \in \mathbf{N}$ and $a_k = 2$ for $k = 0$, $a_k = 1$ otherwise. Therefore, the associated AGT matrix is a proper Riordan Array defined by the triple:

$$d_0 = 1, \quad A = (2, 1, 1, 1, \ldots), \quad Z = (0, 0, 0, 0, \ldots).$$

We have obviously $d(t) = 1$ and since $A(t) = (2 - t)/(1 - t)$ we obtain:

$$h(t) = \frac{1 + t - \sqrt{1 - 6t + t^2}}{2t} = 2 + 2t + 6t^2 + 22t^3 + 90t^4 + 394t^5 + 1806t^6 + 8558t^7 + O(t^8)$$

which is related to the Schröder numbers $S_n = [t^n]\frac{1+t-\sqrt{1-6t+t^2}}{4t}$. By developing around the dominating singularity $t = 3 - 2\sqrt{2}$ and taking the first order approximation, we have:

$$\sum_{k=0}^{n} d_{n,k} = [t^n]\frac{1 - t - \sqrt{1 - 6\,t + t^2}}{2t} \approx -[t^n]\frac{\sqrt{3\,\sqrt{2} - 4}}{3 - 2\,\sqrt{2}}\sqrt{1 - (3 + 2\sqrt{2})t} =$$

$$= \frac{\sqrt{3\,\sqrt{2} - 4}}{(3 - 2\,\sqrt{2})(2n - 1)}\left(\frac{3 + 2\sqrt{2}}{4}\right)^n\binom{2n}{n},$$

$$\sum_{k=0}^{n} k d_{n,k} = [t^n]\frac{1 - 5\,t + 2\,t^2 + (2\,t - 1)\,\sqrt{1 - 6\,t + t^2}}{2t^2} \approx$$

$$\approx -[t^n]\frac{\sqrt{3\sqrt{2}-4}\left(4\sqrt{2}-5\right)}{\left(2\sqrt{2}-3\right)^2}\sqrt{1-(3+2\sqrt{2})t} =$$

$$= \frac{\sqrt{3\sqrt{2}-4}\left(4\sqrt{2}-5\right)}{\left(2\sqrt{2}-3\right)^2(2n-1)}\left(\frac{3+2\sqrt{2}}{4}\right)^n\binom{2n}{n}.$$

Another interesting example is given by the following generating tree specification, which has been used by West in [12]:

$$\left\{\begin{array}{lll} root: & (2) & \\ rule: & (k) & \to (2)^{k-1}(k+1) \end{array}\right. \tag{4.10}$$

In this case we have $c = 2$, $b_k = k - 1$, $\forall k \in \mathbf{N}$ and $a_k = 1$ for $k = 0$, $a_k = 0$ otherwise. The associated AGT matrix is a proper Riordan Array defined as follows:

$$d_0 = 1, \quad A = (1, 0, 0, \ldots), \quad Z = (1, 2, 3, 4, \ldots)$$

This time we have $A(t) = 1$, hence $h(t) = 1$ (this means that the array has equal columns, except for translation). Concerning $d(t)$ we have $Z(t) = 1/(1-t)^2$ and by applying Theorem 3.7 we get:

$$d(t) = \frac{(1-t)^2}{1-3t+t^2}$$

which is the generating function for F_{2n-1}, F_n being the n^{th} Fibonacci number. By developing around the dominating singularity $t = \frac{3-\sqrt{5}}{2}$ and taking the first order approximation we obtain:

$$d_{n,k} = [t^{n-k}]\frac{(1-t)^2}{1-3t+t^2} \approx [t^{n-k}]\frac{1}{\sqrt{5}}\frac{1}{\left(1-\frac{3+\sqrt{5}}{2}t\right)} = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{2(n-k)}.$$

We now conclude illustrating the converse part of Theorem 3.9. In Merlini, Sprugnoli and Verri [5] some classes of lattice paths are studied and their relation with proper Riordan Arrays is shown. Those paths start from the origin and are composed by:

i) *east* steps, i.e., steps from (x, y) to $(x + 1, y)$;

ii) *north-east* (or diagonal) steps, i.e., steps from (x, y) to $(x + 1, y + 1)$;

iii) *north* steps, i.e., steps from (x, y) to $(x, y + 1)$.

Each step can take different colours, and we consider a set of α colours for east steps, a set of β colours for north-east steps and a set of γ colours for north steps. Because of the different colours each step can take on, the paths are called *coloured*. We are interested in *weakly underdiagonal paths*, that is, paths whose ending point (x, y) is such that $x \geq y$; in other words, the ending point is not situated above the main diagonal $x = y$. When the path never goes above this diagonal, we simply call it an *underdiagonal path*.

Shapiro's paths [8] are another interesting kind of paths. These paths are characterized by the fact that the north-east steps on the main diagonal may have different colours from the north-east steps on the other diagonals. In other words, we have (α, β, γ) colours as in the paths above, but on the main diagonal β' colours $(\beta' \neq \beta)$ are used for north-east steps.

The paths made up of n steps and ending at distance k from the main diagonal (the distance is measured along the $x-$ or $y-$axis) can be counted by means of the following theorem, which can be proved by using the results obtained in Merlini, Sprugnoli and Verri [5]. We let $P = \{p_{n,k}\}_{n,k\in\mathbf{N}}$, $Q = \{q_{n,k}\}_{n,k\in\mathbf{N}}$, $S = \{s_{n,k}\}_{n,k\in\mathbf{N}}$ and $R = \{r_{n,k}\}_{n,k\in\mathbf{N}}$ be infinite matrices where:

- $p_{n,k}$ is the number of n-long underdiagonal paths ending at a distance k from the main diagonal;
- $q_{n,k}$ is the number of n-long weakly underdiagonal paths ending at a distance k from the main diagonal;
- $s_{n,k}$ is the number of n-long underdiagonal Shapiro's paths ending at a distance k from the main diagonal;
- $r_{n,k}$ is the number of n-long weakly underdiagonal Shapiro's paths ending at a distance k from the main diagonal,

with colours (α, β, γ), $\alpha \neq 0$, for east, north-east and north steps and β' for the Shapiro's north-east steps on the main diagonal.

The results in Merlini, Sprugnoli and Verri [5] are summarized by the following:

Theorem 4.1 *The infinite arrays* $P = \{p_{n,k}\}_{n,k\in\mathbf{N}}$, $Q = \{q_{n,k}\}_{n,k\in\mathbf{N}}$, $S = \{s_{n,k}\}_{n,k\in\mathbf{N}}$ *and* $R = \{r_{n,k}\}_{n,k\in\mathbf{N}}$ *are proper Riordan Arrays having* $d_0 = 1$, $A = (\alpha, \beta, \gamma, 0, 0, \ldots)$ *as their* A*-sequence and with the following* Z*-sequences:*

$$Z_p = (\beta, \gamma, 0, 0, \ldots), \qquad Z_q = (\beta, 2\gamma, 0, 0, \ldots),$$

$$Z_s = (\beta', \gamma, 0, 0, \ldots), \qquad Z_r = (\beta', 2\gamma, 0, 0, \ldots).$$

We can now use Theorems 3.9 and 4.1 for finding the generating trees specifications which have P, Q, S and R as associated AGT matrices:

Theorem 4.2 *The proper Riordan Arrays* $P, Q, S,$ *and* R *are the* AGT *matrices associated to the following generating tree specifications:*

$$\begin{cases} root: & (c) \\ rule: & (k) \ \rightarrow (c)^{b_k}(k-1)^{\gamma}(k)^{\beta}(k+1)^{\alpha}, \qquad c \in \mathbf{N} \end{cases} \tag{4.11}$$

where:

- *in the* P*-case,* $b_k = 0$, $\forall k \in \mathbf{N}$*;*
- *in the* Q*-case,* $b_{c+1} = \gamma$ *and otherwise* $b_k = 0$*;*
- *in the* S*-case,* $b_c = \beta' - \beta$ *and otherwise* $b_k = 0$*;*
- *in the* R*-case,* $b_c = \beta' - \beta$, $b_{c+1} = \gamma$ *and otherwise* $b_k = 0$.

Acknowledgments

We are greatly indebted to Renzo Sprugnoli for suggesting this research and for his useful comments on the paper.

References

[1] E. Barcucci, A. Del Lungo, E. Pergola, and R. Pinzani. Eco: a general methodology for the Enumeration of Combinatorial Objects. *Advances in combinatorics and differential equations, to appear.*

[2] F. R. K. Chung, R. L. Graham, V. E. Hoggat, and M. Kleiman. The number of Baxter permutations. *Journal of Combinatorial Theory, Series A*, 24:382–394, 1978.

[3] C. Corsani, D. Merlini, and R. Sprugnoli. Left-inversion of combinatorial sums. *Discrete Mathematics*, 180:107–122, 1998.

[4] D. Merlini, D. G. Rogers, R. Sprugnoli, and M. C. Verri. On some alternative characterizations of Riordan arrays. *Canadian Journal of Mathematics*, 49(2):301–320, 1997.

[5] D. Merlini, R. Sprugnoli, and M. C. Verri. Algebraic and combinatorial properties of simple, coloured walks. In *Proceedings of CAAP'94*, volume 787 of *Lecture Notes in Computer Science*, pages 218–233, 1994.

[6] D. G. Rogers. Pascal triangles, Catalan numbers and renewal arrays. *Discrete Mathematics*, 22:301–310, 1978.

[7] G. Rozenberg and A. Salomaa. *The mathematical theory of L systems.* Accademic Press, New York, 1980.

[8] L. W. Shapiro. Private comunication.

[9] L. W. Shapiro, S. Getu, W.-J. Woan, and L. Woodson. The Riordan group. *Discrete Applied Mathematics*, 34:229–239, 1991.

[10] R. Sprugnoli. Riordan arrays and combinatorial sums. *Discrete Mathematics*, 132:267–290, 1994.

[11] J. West. Generating trees and the Catalan and Schröder numbers. *Discrete Mathematics*, 146:247–262, 1995.

[12] J. West. Generating trees and forbidden subsequences. *Discrete Mathematics*, 157:363–374, 1996.