Algebraic aspects of some Riordan arrays related to binary words avoiding a pattern

D. Merlini and R. Sprugnoli
Dipartimento di Sistemi e Informatica
viale Morgagni 65, 50134, Firenze, Italia
merlini@dsi.unifi.it, *renzo.sprugnoli@unifi.it*

17th July 2009

Abstract

We consider some Riordan arrays related to binary words avoiding a pattern which can be easily studied by means of an A-*matrix* rather than their A-*sequence*. Both concepts allow us to define every element as a linear combination of other elements in the array; the A-sequence is unique while the A-matrix is not. However, for the problems considered in the present paper, we show that the A-matrix approach is more convenient. We provide explicit algebraic generating functions for these Riordan arrays and obtain many statistics on the corresponding languages.

1 Introduction

Recently, Baccherini, Merlini and Sprugnoli [1] have studied the relation between binary words excluding a pattern and proper Riordan arrays. In particular, they have proved necessary and sufficient conditions under which the number of words, counted with respect to the number of zeroes and ones, is related to proper Riordan arrays. This problem is interesting in the context of the Riordan arrays theory because the matrices which here arise are naturally defined by recurrence relations following the characterization given in Merlini, Rogers, Sprugnoli and Verri [6]. Some history is necessary at this point.

The concept of a *Riordan array* was introduced in 1991 by Shapiro, Getu, Woan and Woodson [8], with the aim of defining a class of infinite lower triangular arrays with properties analogous to those of the Pascal triangle. This concept was successively studied by Sprugnoli [9] in the context of the computation of combinatorial sums. In these papers, Riordan arrays correspond to matrices where each element $d_{n,k}$ is described by a linear combination of the elements in the previous row, starting from the previous column. The coefficients of this linear combination are independent of n and k and constitute a specific sequence called the A-sequence of the Riordan array. Later, several new characterizations of Riordan arrays were given in [6]: the main result in that paper shows that a lower triangular array $d_{n,k}$ is Riordan whenever its generic element $d_{n+1,k+1}$ linearly depends on the elements $d_{r,s}$ lying in a well-defined, but large zone of the array. The coefficients of this dependence constitute the so-called A-matrix and are illustrated in Figure 1.1. There is no difference between Riordan arrays defined in either way: the A-sequence is a particular case of A-matrix and, given a Riordan array defined by an A-matrix, this corresponds to a well defined A-sequence.

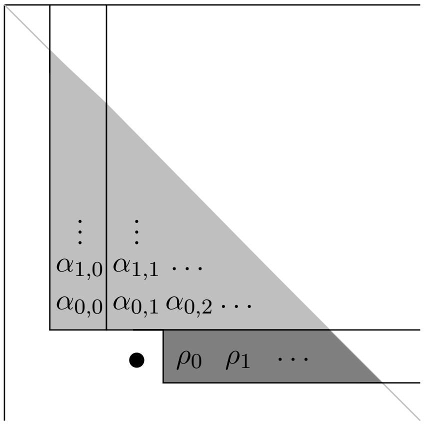

Figure 1.1: The zones which $d_{n+1,k+1}$ can depend on.

However, there are some examples in which a Riordan array can be easily studied by means of the A-matrix while the A-sequence is very complex. This is just the case for the problem studied in [1]. From a combinatorial point of view, this means that it is very challenging to find a construction allowing to obtain objects of size $n+1$ from objects of size n. Instead, the existence of a simple A-matrix corresponds to a possible construction from objects of different sizes less than $n+1$.

The aim of this paper is to re-consider from an algebraic point of view some cases of the problem studied in [1]. In particular, we show that when the pattern to avoid has a particular shape, then the Riordan array under consideration can be explicitly defined in terms of its generating functions. In all these cases, we have a simple A-matrix, while the corresponding A-sequence is (in general) very complicated. We give many examples of languages avoiding some pattern, which can be described by Riordan arrays of this type. The involved generating functions are all algebraic and allow us to find many statistics on the corresponding languages.

2 Riordan arrays for Riordan patterns

In this paper we are interested in studying binary words excluding a given pattern $\mathfrak{p} = p_0 ... p_{h-1} \in \{0,1\}^h$. In particular, if $F^{[\mathfrak{p}]}_{n,k}$ denotes the number of words excluding the pattern and having n bits 1 and k bits 0, then by using the results in Baccherini, Merlini, Sprugnoli [2] we have

$$F^{[\mathfrak{p}]}(x,y) = \sum_{n,k\geq 0} F^{[\mathfrak{p}]}_{n,k} x^n y^k = \frac{C^{[\mathfrak{p}]}(x,y)}{(1-x-y)C^{[\mathfrak{p}]}(x,y) + x^{n_1^{\mathfrak{p}}} y^{n_0^{\mathfrak{p}}}}, \tag{2.1}$$

where $n_1^{[\mathfrak{p}]}$ and $n_0^{[\mathfrak{p}]}$ correspond to the number of ones and zeroes in the pattern and $C^{[\mathfrak{p}]}(x,y)$ is the autocorrelation polynomial with coefficients given by the autocorrelation vector (see also [3, 4, 7]). For a given $\mathfrak{p}$, this vector of bits $c = (c_0, \ldots, c_{h-1})$ can be defined in terms of Iverson's bracket notation (for a predicate P, the expression $[\![P]\!]$ has value 1 if P is true and 0 otherwise) as follows:

$$c_i = [\![p_0 p_1 \cdots p_{h-1-i} = p_i p_{i+1} \cdots p_{h-1}]\!].$$

1	1	0	0	1	1	Tails					
1	1	0	0	1	1						1
	1	1	0	0	1	1					0
		1	1	0	0	1	1				0
			1	1	0	0	1	1			0
				1	1	0	0	1	1		1
					1	1	0	0	1	1	1

Table 2.1: The autocorrelation vector for $\mathfrak{p} = 110011$.

In other words, the bit c_i is determined by shifting $\mathfrak{p}$ right by i positions and setting $c_i = 1$ if and only if the remaining letters match the original. For example, when $\mathfrak{p} = 110011$ the autocorrelation vector is $c = (1, 0, 0, 0, 1, 1)$, as illustrated in Table 2.1, and $C^{[\mathfrak{p}]}(x, y) = 1 + x^2y^2 + x^3y^2$, that is, we mark with x^jy^i the tails of the pattern with j bits 1, i bits 0 and $c_{j+i} = 1$. Therefore, in this case we have:

$$F^{[\mathfrak{p}]}(x, y) = \frac{1 + x^2y^2 + x^3y^2}{(1 - x - y)(1 + x^2y^2 + x^3y^2) + x^4y^2}.$$

n/k	0	1	2	3	4	5	6	7
0	1	1	1	1	1	1	1	1
1	1	2	3	4	5	6	7	8
2	1	3	6	10	15	21	28	36
3	1	4	10	20	35	56	84	120
4	1	5	14	33	67	122	205	324
5	1	6	19	50	114	232	432	750
6	1	7	25	72	181	404	822	1552
7	1	8	32	100	273	660	1451	2952

Table 2.2: The matrix $\mathcal{F}^{[\mathfrak{p}]}$ for $\mathfrak{p} = 110011$

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	20	10	4	1				
4	67	33	14	5	1			
5	232	114	50	19	6	1		
6	822	404	181	72	25	7	1	
7	2952	1451	660	272	100	32	8	1

Table 2.3: The triangle $\mathcal{R}^{[\mathfrak{p}]}$ for $\mathfrak{p} = 110011$

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	20	10	4	1				
4	67	35	15	5	1			
5	232	122	56	21	6	1		
6	822	432	205	84	28	7	1	
7	2952	1552	750	324	120	36	8	1

Table 2.4: The triangle $\mathcal{R}^{[\bar{\mathfrak{p}}]}$ for $\bar{\mathfrak{p}} = 001100$

In order to study the binary words avoiding a pattern in terms of Riordan arrays, we consider the array $\mathcal{R}^{[\mathfrak{p}]} = (R^{[\mathfrak{p}]}_{n,k})$ given by the lower triangular part of the array $\mathcal{F}^{[\mathfrak{p}]} = (F^{[\mathfrak{p}]}_{n,k})$, that is, $R^{[\mathfrak{p}]}_{n,k} = F^{[\mathfrak{p}]}_{n,n-k}$ with $k \leq n$. More precisely, $R^{[\mathfrak{p}]}_{n,k}$ counts the number of words avoiding $\mathfrak{p}$ and having length $2n-k$, n bits one and $n-k$ bits zero. Given a pattern $\mathfrak{p} = p_0 \dots p_{h-1} \in \{0,1\}^h$, let $\bar{\mathfrak{p}} = \bar{p}_0 \dots \bar{p}_{h-1}$ be the pattern with $\bar{p}_i = 1 - p_i, \forall i = 0, \cdots, h-1$. We obviously have $R^{[\bar{\mathfrak{p}}]}_{n,k} = F^{[\bar{\mathfrak{p}}]}_{n,n-k} = F^{[\mathfrak{p}]}_{n-k,n}$, therefore, the matrices $\mathcal{R}^{[\mathfrak{p}]}$ and $\mathcal{R}^{[\bar{\mathfrak{p}}]}$ represent the lower and upper triangular part of the array $\mathcal{F}^{[\mathfrak{p}]}$, respectively. Moreover, we have $R^{[\mathfrak{p}]}_{n,0} = R^{[\bar{\mathfrak{p}}]}_{n,0} = F^{[\mathfrak{p}]}_{n,n}$, $\forall n \in \mathbb{N}$, that is, columns zero of $\mathcal{R}^{[\mathfrak{p}]}$ and $\mathcal{R}^{[\bar{\mathfrak{p}}]}$ correspond to the main diagonal of $\mathcal{F}^{[\mathfrak{p}]}$. Tables 2.2, 2.3 and 2.4 illustrate some rows for the matrices $\mathcal{F}^{[\mathfrak{p}]}$, $\mathcal{R}^{[\mathfrak{p}]}$ and $\mathcal{R}^{[\bar{\mathfrak{p}}]}$ when $\mathfrak{p} = 110011$.

We briefly recall that a Riordan array is an infinite lower triangular array $(d_{n,k})_{n,k\in\mathbb{N}}$, defined by a pair of formal power series $(d(t), h(t))$, such that $d(0) \neq 0, h(0) = 0, h'(0) \neq 0$ and the generic element $d_{n,k}$ is the n-th coefficient in the series $d(t)h(t)^k$, 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 $k > n$. An alternative definition, is in terms of the so-called A-sequence and Z-sequence, with generating functions $A(t)$ and $Z(t)$ satisfying the relations:

$$h(t) = tA(h(t)), \quad d(t) = \frac{d_0}{1 - tZ(h(t))} \quad \text{with} \quad d_0 = d(0).$$

Another characterization states that a lower triangular array $(d_{n,k})_{n,k\in\mathbb{N}}$ is Riordan if and only if there exist another array $(\alpha_{i,j})_{i,j\in\mathbb{N}}$, with $\alpha_{0,0} \neq 0$, and a sequence $(\rho_j)_{j\in\mathbb{N}}$ such that:

$$d_{n+1,k+1} = \sum_{i\geq 0}\sum_{j\geq 0}\alpha_{i,j} d_{n-i,k+j} + \sum_{j\geq 0}\rho_j d_{n+1,k+j+2}. \tag{2.2}$$

Figure 1.1 gives a graphical representation of this kind of dependence. Matrix $(\alpha_{i,j})_{i,j\in\mathbb{N}}$ is called the A-matrix of the Riordan array. If $P^{[0]}(t), P^{[1]}(t), P^{[2]}(t), \dots$ denote the generating functions of rows $0, 1, 2, \dots$ in the A-matrix, i.e.:

$$P^{[i]}(t) = \alpha_{i,0} + \alpha_{i,1}t + \alpha_{i,2}t^2 + \alpha_{i,3}t^3 + \dots$$

and $Q(t)$ is the generating function for the sequence $(\rho_j)_{j\in\mathbb{N}}$, then we have:

$$\frac{h(t)}{t} = \sum_{i\geq 0} t^i P^{[i]}(h(t)) + \frac{h(t)^2}{t} Q(h(t)). \tag{2.3}$$

Another important property of Riordan array concerns the computation of combinatorial sums. In particular we have the following result (see Sprugnoli [9]):

$$\sum_{k=0}^{n} d_{n,k} f_k = [t^n] d(t) f(h(t)) \tag{2.4}$$

that is, every combinatorial sum involving a Riordan array can be computed by extracting the coefficient of t^n from the generating function $d(t)f(h(t))$ where $f(t)$ is the generating function of the sequence $(f_k)_{k\in\mathbb{N}}$. The complete theory of Riordan arrays and the proofs of their properties can be found in [5, 6].

Coming back to our original problem, Baccherini, Merlini, Sprugnoli [2] have proved necessary and sufficient conditions under which the matrices $\mathcal{R}^{[\mathfrak{p}]}$ and $\mathcal{R}^{[\bar{\mathfrak{p}}]}$ are Riordan arrays. In this paper, we are interested to examine into details the case when both $\mathcal{R}^{[\mathfrak{p}]}$ and $\mathcal{R}^{[\bar{\mathfrak{p}}]}$ are Riordan arrays. To this purpose, we introduce the following definition:

Definition 2.1 (Riordan pattern) *We say that* $\mathfrak{p} = p_0 \dots p_{h-1}$ *is a Riordan pattern if and only if*

$$C^{[\mathfrak{p}]}(x,y) = C^{[\mathfrak{p}]}(y,x) = \sum_{i=0}^{\lfloor (h-1)/2 \rfloor} c_{2i} x^i y^i, \quad \textit{and} \quad |n_1^{[\mathfrak{p}]} - n_0^{[\mathfrak{p}]}| \in \{0,1\}\,.$$

For example, the pattern $\mathfrak{p} = 110011$ introduced in Section 2 is not a Riordan pattern. We prove the following result:

Theorem 2.2 *The matrices* $\mathcal{R}^{[\mathfrak{p}]}$ *and* $\mathcal{R}^{[\bar{\mathfrak{p}}]}$ *are both Riordan arrays* $\mathcal{R}^{[\mathfrak{p}]} = (d^{[\mathfrak{p}]}(t), h^{[\mathfrak{p}]}(t))$ *and* $\mathcal{R}^{[\bar{\mathfrak{p}}]} = (d^{[\bar{\mathfrak{p}}]}(t), h^{[\bar{\mathfrak{p}}]}(t))$ *if and only if* $\mathfrak{p}$ *is a Riordan pattern. Moreover we have:*

$$d^{[\mathfrak{p}]}(t) = d^{[\bar{\mathfrak{p}}]}(t) = [x^0] F\left(x, \frac{t}{x}\right) = \frac{1}{2\pi i} \oint F\left(x, \frac{t}{x}\right) \frac{dx}{x}$$

and

$$h^{[\mathfrak{p}]}(t) = \frac{1 - \sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,1} t^{i+1} - \sqrt{(1 - \sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,1} t^{i+1})^2 - 4 \sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,0} t^{i+1} (\sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,2} t^{i+1} + 1)}}{2(\sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,2} t^{i+1} + 1)}$$

where $\delta_{i,j}$ *is the Kronecker delta,*

$$\sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,0} t^i = \sum_{i=0}^{n_1^{\mathfrak{p}}-1} c_{2i} t^i - \delta_{-1, n_0^{\mathfrak{p}} - n_1^{\mathfrak{p}}} t^{n_1^{\mathfrak{p}}-1},$$

$$\sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,1} t^i = -\sum_{i=0}^{n_1^{\mathfrak{p}}-1} c_{2(i+1)} t^i - \delta_{0, n_0^{\mathfrak{p}} - n_1^{\mathfrak{p}}} t^{n_1^{\mathfrak{p}}-1},$$

$$\sum_{i=0}^{n_1^{\mathfrak{p}}-1} \alpha_{i,2} t^i = \sum_{i=0}^{n_1^{\mathfrak{p}}-1} c_{2(i+1)} t^i - \delta_{1, n_0^{\mathfrak{p}} - n_1^{\mathfrak{p}}} t^{n_1^{\mathfrak{p}}-1},$$

and the coefficients c_i *are given by the autocorrelation vector of* $\mathfrak{p}$. *An analogous formula holds for* $h^{[\bar{\mathfrak{p}}]}(t)$.

Proof: The proof follows directly from [1, Theorem 3.1] and consists in extracting the coefficients

$$[x^{n+1}y^{k+1}]\left((1-x-y)C^{[\mathfrak{p}]}(x,y)+x^{n_1^{\mathfrak{p}}}y^{n_0^{\mathfrak{p}}}\right)F^{[\mathfrak{p}]}(x,y)=[x^{n+1}y^{k+1}]C^{[\mathfrak{p}]}(x,y)$$

and then putting $R^{[\mathfrak{p}]}_{n,k}=F^{[\mathfrak{p}]}_{n,n-k}=[x^n y^{n-k}]F^{[\mathfrak{p}]}(x,y)$. We thus obtain:

$$R^{[\mathfrak{p}]}_{n+1,k+1}=R^{[\mathfrak{p}]}_{n,k}+R^{[\mathfrak{p}]}_{n+1,k+2}-R^{[\mathfrak{p}]}_{n+1-n_1^{\mathfrak{p}},k+1+n_0^{\mathfrak{p}}-n_1^{\mathfrak{p}}}+$$

$$-\sum_{i\geq 1}c_{2i}\left(R^{[\mathfrak{p}]}_{n+1-i,k+1}-R^{[\mathfrak{p}]}_{n-i,k}-R^{[\mathfrak{p}]}_{n+1-i,k+2}\right)$$

and the formula for $h^{[\mathfrak{p}]}(t)$ follows from equation (2.3) with $P^{[i]}(t)=\alpha_{i,0}+\alpha_{i,1}t+\alpha_{i,2}t^2$ and $Q(t)=1$. For what concerns $d^{[\mathfrak{p}]}(t)$, we simply use the Cauchy formula for finding the main diagonal of matrix $\mathcal{F}^{[\mathfrak{p}]}$ (see, e.g., [10, Chap. 6, p. 182]). ∎

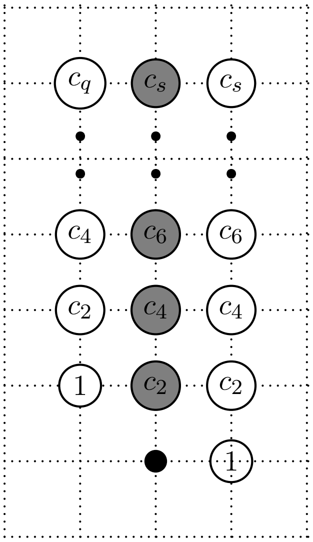

Figure 2.2: The A-matrix corresponding to a Riordan pattern: the coefficients in the gray circles are negative, $s=2n_1^{\mathfrak{p}}$, $q=2(n_1^{\mathfrak{p}}-1)$. Moreover, we have to consider the contribution of $-R^{[\mathfrak{p}]}_{n+1-n_1^{\mathfrak{p}},k+1+n_0^{\mathfrak{p}}-n_1^{\mathfrak{p}}}$.

As a direct consequence of the previous theorem we have:

Corollary 2.3 *Let $\mathfrak{p}$ be a Riordan pattern. Then the Riordan array $\mathcal{R}^{[\mathfrak{p}]}$ is characterized by the A-matrix defined by the following relation:*

$$R^{[\mathfrak{p}]}_{n+1,k+1}=R^{[\mathfrak{p}]}_{n,k}+R^{[\mathfrak{p}]}_{n+1,k+2}-R^{[\mathfrak{p}]}_{n+1-n_1^{\mathfrak{p}},k+1+n_0^{\mathfrak{p}}-n_1^{\mathfrak{p}}}+$$

$$-\sum_{i\geq 1}c_{2i}\left(R^{[\mathfrak{p}]}_{n+1-i,k+1}-R^{[\mathfrak{p}]}_{n-i,k}-R^{[\mathfrak{p}]}_{n+1-i,k+2}\right),$$

where the coefficients c_i are given by the autocorrelation vector of $\mathfrak{p}$.

Figure 2.2, gives a graphical representation of the A-matrix: in particular, $s = 2n_1^{\mathfrak{p}}$, $q = 2(n_1^{\mathfrak{p}}-1)$ and the coefficients in the gray circles have to be taken as negative while the coefficients in the white circles have to be taken as positive. Moreover we have to consider the contribution of $-R^{[\mathfrak{p}]}_{n+1-n_1^{\mathfrak{p}},k+1+n_0^{\mathfrak{p}}-n_1^{\mathfrak{p}}}$.

By specializing Theorem 2.2 to the cases $|n_1^{\mathfrak{p}}-n_0^{\mathfrak{p}}| \in \{0,1\}$ and setting $C^{[\mathfrak{p}]}(t) = C^{[\mathfrak{p}]}(\sqrt{t},\sqrt{t}) = \sum_{i\geq 0} c_{2i}t^i$, we have the following corollaries:

Corollary 2.4 *Let $\mathfrak{p}$ be a Riordan pattern with $n_1^{\mathfrak{p}} - n_0^{\mathfrak{p}} = 1$. Then we have:*

$$d^{[\mathfrak{p}]}(t) = \frac{C^{[\mathfrak{p}]}(t)}{\sqrt{C^{[\mathfrak{p}]}(t)^2 - 4tC^{[\mathfrak{p}]}(t)(C^{[\mathfrak{p}]}(t) - t^{n_0^{\mathfrak{p}}})}},$$

$$h^{[\mathfrak{p}]}(t) = \frac{C^{[\mathfrak{p}]}(t) - \sqrt{C^{[\mathfrak{p}]}(t)^2 - 4tC^{[\mathfrak{p}]}(t)(C^{[\mathfrak{p}]}(t) - t^{n_0^{\mathfrak{p}}})}}{2C^{[\mathfrak{p}]}(t)}.$$

Proof: From Theorem 2.2 we have

$$d^{[\mathfrak{p}]}(t) = [x^0]F^{[\mathfrak{p}]}\left(x, \frac{t}{x}\right) = \frac{1}{2\pi i}\oint F^{[\mathfrak{p}]}\left(x, \frac{t}{x}\right)\frac{dx}{x}$$

and when $n_1^{\mathfrak{p}} - n_0^{\mathfrak{p}} = 1$ we obtain:

$$\frac{1}{x}F^{[\mathfrak{p}]}\left(x, \frac{t}{x}\right) = \frac{-C^{[\mathfrak{p}]}(t)}{x^2(C^{[\mathfrak{p}]}(t) - t^{n_0^{\mathfrak{p}}}) - xC^{[\mathfrak{p}]}(t) + tC^{[\mathfrak{p}]}(t)}.$$

In order to compute the integral, it is necessary to find the singularities $x(t)$ such that $x(t) \to 0$ with $t \to 0$ and apply the Residue theorem. We have two singularities ($x_1(t)$ corresponds to the plus sign and $x_2(t)$ to the minus):

$$x_{1,2}(t) = \frac{C^{[\mathfrak{p}]}(t) \pm \sqrt{C^{[\mathfrak{p}]}(t)^2 - 4tC^{[\mathfrak{p}]}(t)(C^{[\mathfrak{p}]}(t) - t^{n_0^{\mathfrak{p}}})}}{2(C^{[\mathfrak{p}]}(t) - t^{n_0^{\mathfrak{p}}})}$$

and

$$\frac{1}{x}F^{[\mathfrak{p}]}\left(x, \frac{t}{x}\right) = \frac{-C^{[\mathfrak{p}]}(t)}{(C^{[\mathfrak{p}]}(t) - t^{n_0^{\mathfrak{p}}})(x - x_1(t))(x - x_2(t))}.$$

Therefore, we have:

$$d^{[\mathfrak{p}]}(t) = \lim_{x\to x_2(t)} \frac{1}{x}F^{[\mathfrak{p}]}\left(x, \frac{t}{x}\right)(x - x_2(t)) = \frac{C^{[\mathfrak{p}]}(t)}{(C^{[\mathfrak{p}]}(t) - t^{n_0^{\mathfrak{p}}})(x_1(t) - x_2(t))}$$

and, after some simplification, we obtain the formula in the statement. The expression for $h^{[\mathfrak{p}]}(t)$ follows directly from Theorem 2.2, by specializing the Kronecker deltas. ∎

For example, when $\mathfrak{p} = 11100$ we have the Riordan array defined by the following functions:

$$d^{[\mathfrak{p}]}(t) = \frac{1}{\sqrt{1-4t+4t^3}}, \quad h^{[\mathfrak{p}]}(t) = \frac{1-\sqrt{1-4t+4t^3}}{2}$$

and illustrated in Table (2.5). As another example, the triangle in Table (2.6) corresponds to

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	18	9	4	1				
4	58	29	13	5	1			
5	192	96	44	18	6	1		
6	650	325	151	64	24	7	1	
7	2232	1116	524	288	90	31	8	1

Table 2.5: The triangle for $\mathfrak{p} = 11100$

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	4	2	1					
3	10	5	2	1				
4	26	13	6	2	1			
5	70	35	16	7	2	1		
6	192	96	45	19	8	2	1	
7	534	267	126	56	22	9	2	1

Table 2.6: The triangle for $\mathfrak{p} = 101$

$\mathfrak{p} = 101$ and to the functions:

$$d^{[\mathfrak{p}]}(t) = \frac{1+t}{\sqrt{1-2t-3t^2}}, \quad h^{[\mathfrak{p}]}(t) = \frac{1+t-\sqrt{1-2t-3t^2}}{2(1+t)}.$$

We observe that in Tables 2.5 and 2.6 every element in column 0, except the first one, is twice the value in column 1. In fact we have the following result:

Theorem 2.5 *Let $\mathfrak{p}$ be a Riordan pattern with $n_1^{\mathfrak{p}} - n_0^{\mathfrak{p}} = 1$. Then the Riordan array $\mathcal{R}^{[\mathfrak{p}]}$ satisfies the following relation:*

$$R^{[\mathfrak{p}]}_{n+1,0} = 2R^{[\mathfrak{p}]}_{n+1,1}.$$

Proof: The relation can be found by observing that by Corollary 2.4 we have

$$\frac{d^{[\mathfrak{p}]}(t)-1}{h^{[\mathfrak{p}]}(t)} = 2.$$

■

The proof of Corollaries 2.6 and 2.7 can be done similarly to Corollary 2.4.

Corollary 2.6 *Let $\mathfrak{p}$ be a Riordan pattern with $n_1^{\mathfrak{p}} - n_0^{\mathfrak{p}} = 0$. Then we have:*

$$d^{[\mathfrak{p}]}(t) = \frac{C^{[\mathfrak{p}]}(t)}{\sqrt{(C^{[\mathfrak{p}]}(t) + t^{n_0^{\mathfrak{p}}})^2 - 4tC^{[\mathfrak{p}]}(t)^2}},$$

$$h^{[\mathfrak{p}]}(t) = \frac{C^{[\mathfrak{p}]}(t) + t^{n_0^{\mathfrak{p}}} - \sqrt{(C^{[\mathfrak{p}]}(t) + t^{n_0^{\mathfrak{p}}})^2 - 4tC^{[\mathfrak{p}]}(t)^2}}{2C^{[\mathfrak{p}]}(t)}.$$

Corollary 2.7 *Let $\mathfrak{p}$ be a Riordan pattern with $n_0^{\mathfrak{p}} - n_1^{\mathfrak{p}} = 1$. Then we have:*

$$d^{[\mathfrak{p}]}(t) = \frac{C^{[\mathfrak{p}]}(t)}{\sqrt{C^{[\mathfrak{p}]}(t)^2 - 4tC^{[\mathfrak{p}]}(t)(C^{[\mathfrak{p}]}(t) - t^{n_1^{\mathfrak{p}}})}},$$

$$h^{[\mathfrak{p}]}(t) = \frac{C^{[\mathfrak{p}]}(t) - \sqrt{C^{[\mathfrak{p}]}(t)^2 - 4tC^{[\mathfrak{p}]}(t)(C^{[\mathfrak{p}]}(t) - t^{n_1^{\mathfrak{p}}})}}{2(C^{[\mathfrak{p}]}(t) - t^{n_1^{\mathfrak{p}}})}.$$

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	5	3	1					
3	15	8	4	1				
4	46	25	12	5	1			
5	147	79	39	17	6	1		
6	477	256	128	58	23	7	1	
7	1570	841	424	198	83	30	8	1

Table 2.7: The triangle for $\mathfrak{p} = 0101$

For the sake of completeness, the triangle in Table (2.7) corresponds to the pattern $\mathfrak{p} = 0101$ and to the functions:

$$d^{[\mathfrak{p}]}(t) = \frac{1+t}{\sqrt{1-2t-5t^2-2t^3+t^4}}, \quad h^{[\mathfrak{p}]}(t) = \frac{1+t+t^2-\sqrt{1-2t-5t^2-2t^3+t^4}}{2(1+t)}.$$

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	18	10	4	1				
4	58	32	15	5	1			
5	192	106	52	21	6	1		
6	650	357	180	79	28	7	1	
7	2232	1222	624	288	114	36	8	1

Table 2.8: The triangle for $\mathfrak{p} = 00011$

Finally, for $\mathfrak{p} = 00011$ we have the triangle illustrated in Table (2.8) and the functions:

$$d^{[\mathfrak{p}]}(t) = \frac{1}{\sqrt{1-4t+4t^3}}, \quad h^{[\mathfrak{p}]}(t) = \frac{1-\sqrt{1-4t+4t^3}}{2(1-t^2)}.$$

As we pointed out in the Introduction, the Riordan arrays studied in this paper are characterized by a complex A-sequence. In fact, the generating function $A(t)$ of this sequence can be found from the relation $h(t) = tA(h(t))$, and, due to Theorem 2.2, we can expect, in general, a complex solution. For example, for the Riordan array in Table 2.8, corresponding to $\mathfrak{p} = 00011$, we find after setting $R(t) = \sqrt{1 + 4t^4 - 4t^3}$:

$$A(t) = \frac{\left(2t^3 - t^2 - t - 1 - (t^2 + t + 1)R(t)\right)\left(2t^3 - \sqrt{2}\sqrt{2t^6 + 8t^4 - 12t^3 + 4 - (4 - 4t^3)R(t)}\right)}{8t^4(t-1)(t+1)}$$

$$= 1+t+t^2+t^4+t^5+2t^7+t^8-t^9+5t^{10}-t^{11}-4t^{12}+16t^{13}-14t^{14}-8t^{15}+57t^{16}-83t^{17}+15t^{18}+197t^{19}+O(t^{20}).$$

Fortunately, for Riordan patterns, we have always a simple A-matrix, as we proved in Corollary 2.3. For example, in the case $\mathfrak{p} = 00011$, we have $R^{[\mathfrak{p}]}_{n+1,k+1} = R^{[\mathfrak{p}]}_{n,k} + R^{[\mathfrak{p}]}_{n+1,k+2} - R^{[\mathfrak{p}]}_{n-2,k}$.

Up to this moment, we gave specific examples. In some cases, we can find general formulas relative to whole classes of patterns:

Theorem 2.8 *For $\mathfrak{p} = 1^{j+1}0^j$ we have the Riordan array:*

$$d^{[\mathfrak{p}]}(t) = \frac{1}{\sqrt{1 - 4t + 4t^{j+1}}}, \quad h^{[\mathfrak{p}]}(t) = \frac{1 - \sqrt{1 - 4t + 4t^{j+1}}}{2};$$

for $\mathfrak{p} = 0^{j+1}1^j$ we have the Riordan array:

$$d^{[\mathfrak{p}]}(t) = \frac{1}{\sqrt{1 - 4t + 4t^{j+1}}}, \quad h^{[\mathfrak{p}]}(t) = \frac{1 - \sqrt{1 - 4t + 4t^{j+1}}}{2(1 - t^j)};$$

for $\mathfrak{p} = 1^j0^j$ we have the Riordan array:

$$d^{[\mathfrak{p}]}(t) = \frac{1}{\sqrt{1 - 4t + 2t^j + t^{2j}}}, \quad h^{[\mathfrak{p}]}(t) = \frac{1 + t^j - \sqrt{1 - 4t + 2t^j + t^{2j}}}{2};$$

for $\mathfrak{p} = (10)^j1$ we have the Riordan array:

$$d^{[\mathfrak{p}]}(t) = \frac{\sum_{i=0}^{j} t^i}{\sqrt{1 - 2\sum_{i=0}^{j} t^i - 3\left(\sum_{i=0}^{j} t^i\right)^2}}, \quad h^{[\mathfrak{p}]}(t) = \frac{\sum_{i=0}^{j} t^i - \sqrt{1 - 2\sum_{i=0}^{j} t^i - 3\left(\sum_{i=0}^{j} t^i\right)^2}}{2\sum_{i=0}^{j} t^i}.$$

We conclude this extended abstract, by observing that by formula (2.4) and Corollaries 2.4, 2.6 and 2.7 we can compute many statistics on the languages considered in the present paper. For example we have for $f_k = 1$ and $f_k = k$:

$$\sum_{k=0}^{n} R^{[\mathfrak{p}]}_{n,k} = [t^n]\frac{d^{[\mathfrak{p}]}(t)}{1 - h^{[\mathfrak{p}]}(t)},$$

$$\sum_{k=0}^{n} kR^{[\mathfrak{p}]}_{n,k} = [t^n]\frac{d^{[\mathfrak{p}]}(t)h^{[\mathfrak{p}]}(t)}{(1 - h^{[\mathfrak{p}]}(t))^2}.$$

In some cases these coefficients can be extracted as exact formulas; in other cases, we can compute an asymptotic approximation of the coefficients.

The problems considered in this paper have a natural combinatorial interpretation in terms of lattice paths where a 1 corresponds to a north-east step and a 0 to a south-east step. These paths start from the origin and avoid the sub-path corresponding to the pattern. For example, let us consider the case $\mathfrak{p} = 1^3 0^2$. In the lattice path interpretation $R^{[\mathfrak{p}]}_{n,k}$ counts the number of paths avoiding $\mathfrak{p}$, having length $2n-k$ and with n north-east steps. Then, by using formula (2.4) and the first case of Theorem 2.8 with $j=2$, we can compute the following statistic:

$$\sum_{k=0}^{n} R^{[\mathfrak{p}]}_{n,k} 2^{n-k} = 2^n [t^n] \frac{d^{[\mathfrak{p}]}(t)}{1-\frac{1}{2}h^{[\mathfrak{p}]}(t)} =$$

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

$$= 2^n [t^n] \left(1 + \frac{5}{2}t + \frac{31}{4}t^2 + \frac{189}{8}t^3 + \frac{1223}{16}t^4 + \frac{8117}{32}t^5 + O(t^6)\right)$$

This statistic has a nice combinatorial interpretation: it counts the number of paths, avoiding the sub-path $\mathfrak{p} = 1^3 0^2$, of length between n and $2n$, having n north-east steps and having the south-east steps of two different colours. With the help of *Maple*, we can find an asymptotic approximation of the coefficients by observing that the polynomial $1-4t+4t^3$ has the following three real roots:

$$x_1 = \frac{2}{3}\sqrt{3}s \sim 0.83756, \quad x_2 = -\frac{1}{3}\sqrt{3}s - q \sim -1.10715, \quad x_3 = -\frac{1}{3}\sqrt{3}s + q \sim 0.26959$$

where

$$s = \cos\left(-\frac{1}{3}\arctan\left(\frac{\sqrt{3}\sqrt{37}}{9}\right) + \frac{\pi}{3}\right), \quad q = \sin\left(-\frac{1}{3}\arctan\left(\frac{\sqrt{3}\sqrt{37}}{9}\right) + \frac{\pi}{3}\right).$$

Therefore, by developing the above generating function around its minimal modulo singularity x_3 and taking the first term, we have:

$$\sum_{k=0}^{n} R^{[\mathfrak{p}]}_{n,k} 2^{n-k} \approx 2^n [t^n] \frac{4\sqrt[4]{3}}{9\sqrt{\frac{(\sqrt{3}s-q)q}{s(\sqrt{3}s+3q)}}} \frac{1}{\sqrt{\left(1-\frac{t}{x_3}\right)}} =$$

$$= \frac{4\sqrt[4]{3}}{9\sqrt{\frac{(\sqrt{3}s-q)q}{s(\sqrt{3}s+3q)}}} \frac{1}{2^n}\binom{2n}{n}\left(\frac{1}{x_3}\right)^n = 1.45198\binom{2n}{n}(1.85463)^n.$$

For example, for $n = 50$ the exact value is $0.3793008365 \cdot 10^{43}$ against an approximate value of $0.3791140086 \cdot 10^{43}$, with a relative error of the 0.05%.

Many other statistics on the languages on $\{0, 1\}$ avoiding a Riordan pattern can be found similarly.

References

[1] D. Baccherini, D. Merlini, and R. Sprugnoli. Binary words excluding a pattern and proper Riordan arrays. *Discrete Mathematics*, 307:1021–1037, 2007.

[2] D. Baccherini, D. Merlini, and R. Sprugnoli. Level Generating Trees and proprer Riordan arrays. *Applicable Analysis and Discrete Mathematics*, 2(1):69–91, 2008.

[3] L. J. Guibas and M. Odlyzko. Long repetitive patterns in random sequences. *Zeitschrift für Wahrscheinlichkeitstheorie*, 53:241–262, 1980.

[4] L. J. Guibas and M. Odlyzko. String overlaps, pattern matching, and nontransitive games. *Journal of Combinatorial Theory, Series A*, 30:183–208, 1981.

[5] D. Merlini. I Riordan Array nell'Analisi degli Algoritmi. Tesi di Dottorato, Università degli Studi di Firenze, 1996.

[6] 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.

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

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

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

[10] R. P. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, Cambridge, 1999.