Automata on Patterns and Graphs

Symeon Bozapalidis and Antonios Kalampakas

Aristotle University of Thessaloniki,
Department of Mathematics,
54124, Thessaloniki, Greece
e-mail:{bozapali,akalamp}@math.auth.gr

Abstract

Magmoids satisfying the 15 fundamental equations of graphs, namely D-magmoids, are introduced. Automata on directed hypergraphs are defined by virtue of finite relational D-magmoids. Two different modes of graph recognizability arise, their closure properties are investigated, and a comparison is being made between the two classes.

1 Introduction

A hypergraph consists of a set of nodes and a set of hyperedges, just as an ordinary (directed) graph except that a hyperedge may have an arbitrary sequence of sources and an arbitrary sequence of targets. Each hyperedge is labelled with a symbol from a doubly ranked alphabet Σ in such a way that the first (second) rank of its label equals the number of its sources (targets respectively). Also, every hypergraph is multi-pointed in the sense that it has a sequence of m "begin" and n "end" nodes, $m, n \geq 0$.

From now on a hypergraph will also be called a graph, and its hyperedges edges; furthermore, to specify the number of begin and end nodes, it will be called an (m, n)-graph. We denote by $GR_{m,n}(\Sigma)$ the set of all (m, n)-graphs labelled over Σ. Any graph $G \in GR_{m,n}(\Sigma)$ having no edges, is called discrete.

If G is an (m, n)-graph and H is an (n, k)-graph then their *product* $G \circ H$ is the (m, k)-graph obtained by taking the disjoint union of G and H and then identifying the ith end node of G with the ith begin node of H, for every $i \in \{1, ..., n\}$; also, the sequence of begin nodes of $G \circ H$ is the one of G, and its sequence of end nodes the one of H.

The *sum* $G \square H$ of arbitrary graphs G and H is their disjoint union with their sequences of begin nodes concatenated and similarly for their end nodes.

The family $GR(\Sigma) = (GR_{m,n}(\Sigma))_{m,n\in\mathbb{N}}$ with the operations $\circ$ and $\square$ forms a magmoid in the sense of [1, 2], that is, a strict monoidal category (or x-category) whose objects are the natural numbers (see e.g. [11, 8]). The algebraic structure of magmoids seems to be the suitable framework for representing and generating directed labelled hypergraphs primarily due to two critical advantages: its generative power and its finite axiomatization.

Indeed, Engelfriet and Vereijken proved that, $GR(\Sigma)$ is finitely generated, that is, any graph can be built from a specific finite set of elementary graphs (cf. [6]). More precisely, let us denote by $I_{p,q}$ the discrete (p, q)-graph having a single node x and whose begin and end sequences are $x \cdots x$ (p times) and $x \cdots x$ (q times) respectively. Let also Π be the discrete $(2, 2)$-graph having two nodes x and y and whose begin and end sequences are xy and yx, respectively, and for $n \geq 0$, let E_n be the graph with n nodes $x_1, \ldots, x_n$ whose begin and end sequence is $x_1 \cdots x_n$; we also write E for E_1. Note that the graphs E_n are the units in the category $GR(\Sigma)$, i.e., if G is an (m, n)-graph, then $E_m \circ G = G$ and $G \circ E_n = G$. Finally, for every $\sigma \in \Sigma_{m,n}$, we denote by $G(\sigma)$ the (m, n)-graph having only one edge and $m + n$ nodes $x_1, \ldots, x_m, y_1, \ldots, y_n$. The edge is labelled by σ, and the begin (resp. end sequence) of the graph is the sequence of sources (resp. targets) of the edge, viz. $x_1 \cdots x_m$ (resp. $y_1 \cdots y_n$).

The following important result is due to Engelfriet and Vereijken (see Theorem 7 of [6]).

Theorem. *Any (m, n)-graph over the finite, doubly ranked alphabet Σ can be constructed from the graphs of the set*

$$\{G(\sigma) \mid \sigma \in \Sigma\} \cup \{I_{2,1}, I_{0,1}, I_{1,2}, I_{1,0}, \Pi\}$$

by using the operations $\circ$ and $\Box$, and the unit graphs E_n.

Note that not all units are needed (in fact just the empty graph E_0) because $E = I_{1,1} = I_{1,2} \circ I_{2,1}$ and $E_n = E \Box \cdots \Box E$ (n times).

Now let us introduce the alphabet D, formed by the following five symbols

$$i_{21} : 2 \to 1 \quad i_{01} : 0 \to 1 \quad i_{12} : 1 \to 2 \quad i_{10} : 1 \to 0 \quad \pi : 2 \to 2$$

where $x : m \to n$ indicates that symbol x has first rank m and second rank n, and denote by $mag(X_\Sigma)$ the free magmoid generated by the doubly ranked alphabet $X_\Sigma = \Sigma \cup D$. The set $mag(X_\Sigma)$ consists of all expressions built from the constants in X_Σ, the binary operations $\circ$ and $\Box$, and the constants e_n ($n \geq 0$; the units of the magmoid), modulo the laws of magmoids. We call the elements of $mag(X_\Sigma)$ *patterns* over X_Σ.

In what follows, the operations $\circ$ and $\Box$ of $mag(X_\Sigma)$ will also be denoted as horizontal and vertical concatenation:

$$\alpha \circ \beta \text{ as } \alpha\beta \text{ , and } \alpha \Box \beta \text{ as } \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

respectively. We denote by

$$val_\Sigma : mag(X_\Sigma) \to GR(\Sigma)$$

the unique magmoid morphism extending the function described by the assignments

$$i_{21} \mapsto I_{2,1}, \quad i_{01} \mapsto I_{0,1}, \quad i_{12} \mapsto I_{1,2}, \quad i_{10} \mapsto I_{1,0}, \quad \pi \mapsto \Pi,$$

$$\sigma \mapsto G(\sigma), \text{ for all } \sigma \in \Sigma, \qquad e_n \mapsto E_n, \text{ for all } n \in \mathbb{N}.$$

The previous theorem implies that the morphism val_Σ is a surjection. However, val_Σ is not an injection and in fact, for any given hypergraph, there are infinitely many patterns representing it.

This ambiguity was recently settled by constructing a *finite* set of equations with the property that two patterns represent the same hypergraph if and only if one can be transformed into the other through these equations (cf. [4]).

More precisely, we denote by $\pi_{m,n}$ the pattern inductively defined by

- $\pi_{1,0} = e$, $\pi_{1,n} = \begin{pmatrix} e_{n-1} \\ \pi \end{pmatrix} \begin{pmatrix} \pi_{1,n-1} \\ e \end{pmatrix}$. Notice that for $n = 1$, $\pi_{1,1} = \pi$.
- $\pi_{0,n} = e_n$, $\pi_{m,n} = \begin{pmatrix} \pi_{m-1,n} \\ e \end{pmatrix} \begin{pmatrix} e_{m-1} \\ \pi_{1,n} \end{pmatrix}$.

Given a finite doubly ranked alphabet Σ, the set of equations

$$\mathcal{E}: \quad \pi\pi = e_2, \quad \begin{pmatrix} e \\ \pi \end{pmatrix}\begin{pmatrix} \pi \\ e \end{pmatrix}\begin{pmatrix} e \\ \pi \end{pmatrix} = \begin{pmatrix} \pi \\ e \end{pmatrix}\begin{pmatrix} e \\ \pi \end{pmatrix}\begin{pmatrix} \pi \\ e \end{pmatrix}, \quad \begin{pmatrix} e \\ i_{21} \end{pmatrix} i_{21} = \begin{pmatrix} i_{21} \\ e \end{pmatrix} i_{21},$$

$$\begin{pmatrix} e \\ i_{01} \end{pmatrix} i_{21} = e, \;\; \pi i_{21} = i_{21}, \quad \begin{pmatrix} \pi \\ e \end{pmatrix}\begin{pmatrix} e \\ \pi \end{pmatrix}\begin{pmatrix} i_{21} \\ e \end{pmatrix} = \begin{pmatrix} e \\ i_{21} \end{pmatrix}\pi, \quad \begin{pmatrix} e \\ i_{01} \end{pmatrix}\pi = \begin{pmatrix} i_{01} \\ e \end{pmatrix},$$

$$i_{12}\begin{pmatrix} e \\ i_{12} \end{pmatrix} = i_{12}\begin{pmatrix} i_{12} \\ e \end{pmatrix}, \quad i_{12}\begin{pmatrix} e \\ i_{10} \end{pmatrix} = e, \quad i_{12}\pi = i_{12}, \quad i_{12}\, i_{21} = e,$$

$$\begin{pmatrix} i_{12} \\ e \end{pmatrix}\begin{pmatrix} e \\ \pi \end{pmatrix}\begin{pmatrix} \pi \\ e \end{pmatrix} = \pi\begin{pmatrix} e \\ i_{12} \end{pmatrix}, \quad \pi\begin{pmatrix} e \\ i_{10} \end{pmatrix} = \begin{pmatrix} i_{10} \\ e \end{pmatrix}, \quad \begin{pmatrix} i_{12} \\ e \end{pmatrix}\begin{pmatrix} e \\ i_{21} \end{pmatrix} = i_{21}\, i_{12},$$

$$\pi_{p,1}\begin{pmatrix} \sigma \\ e \end{pmatrix} = \begin{pmatrix} e \\ \sigma \end{pmatrix}\pi_{q,1}, \quad \text{where } \sigma \in \Sigma_{p,q}, \;\; p, q \geq 0,$$

has the following property: for all patterns p and q,

$$val_\Sigma(p) = val_\Sigma(q) \text{ if and only if } p \underset{\mathcal{E}}{=} q.$$

Therefore, $GR(\Sigma)$ is characterized as the quotient of the free magmoid generated by Σ, divided by $\mathcal{E}$, or equivalently, it is the free object generated by Σ within the category of all magmoids $M = (M_{m,n})$, which are endowed with elements $\pi \in M_{2,2}$, $i_{21} \in M_{2,1}$, $i_{01} \in M_{0,1}$, $i_{12} \in M_{1,2}$, $i_{10} \in M_{1,0}$, satisfying the equations $\mathcal{E}$.

In this respect, various algebraic properties of graphs and graph languages can be investigated inside the framework of magmoids. Our aim in the present paper is to study automata on patterns and graphs.

The paper is divided into 8 sections. The notion of a magmoid, together with some preliminary matter, is presented in Section 2. Examples of this algebraic structure are considered: magmoids of functions and magmoids of relations. We particularly insist in the construction of the *magmoid of hypergraphs* by recalling the definition of hypergraphs introduced in [6] together with the operations product and sum. In Section 3, we construct the free magmoid generated by a doubly ranked alphabet and give the definitions of pattern and pattern language.

The deterministic and nondeterministic pattern automata are presented in Section 4. Our nondeterministic pattern automaton was first introduced in [3], where a Kleene Theorem is presented for these devices. We prove that the deterministic and nondeterministic classes are equivalent and prove that this class is closed under $\Box$, $\Box$-star, intersection, magmoid morphisms and inverse alphabetic morphisms.

In Section 5 we introduce the D-magmoid, i.e., a magmoid equipped with five elements satisfying the equations $\mathcal{E}$. A D-magmoid over the magmoid of relations is called relational. Two important examples of relational D-magmoids are presented: the diagonal and the group D-magmoid. In the second case the D-magmoid is constructed by virtue of an arbitrary commutative group. We are particulary interested in group D-magmoids corresponding to the cyclic groups $\mathcal{Z}_m$, $m \geq 2$. Moreover, we prove that the set of all graphs over the doubly ranked alphabet Σ, is the free D-magmoid generated by Σ.

Automata on graphs are introduced in Section 6. Graph automata are defined with respect to a finite relational D-magmoid. Two different modes of graph recognizability arise, namely the diagonal and the group recognizability, corresponding respectively to the diagonal and the group D-magmoid.

Closure properties of the two classes are investigated in Section 7. Diagonal and group recognizability is proved to be closed under union, intersection and graph homomorphism. Moreover, the inverse image of a (diagonal or group) recognizable graph language via the magmoid morphism val_Σ is a recognizable pattern language.

In the last Section we investigate the hierarchy inside the class of $\mathcal{Z}_m$-recognizable graph languages, $m \geq 2$. We prove that $m \mid n$ if and only if every $\mathcal{Z}_m$-recognizable graph language is $\mathcal{Z}_n$-recognizable. Moreover, we prove that the classes of diagonal and group recognizable graph languages are incomparable but not disjoint.

2 Magmoids

Recall that a doubly ranked set (or a doubly ranked alphabet) $(A_{m,n})_{m,n\in\mathbb{N}}$ is a set A together with a function $rank : A \to \mathbb{N} \times \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers. For $m, n \in \mathbb{N}$, $A_{m,n}$ is the set $\{a \in A \mid rank(a) = (m, n)\}$. In what follows we will drop the subscript $m, n \in \mathbb{N}$ and denote a doubly ranked set simply by $(A_{m,n})$.

A *magmoid* is a doubly ranked set $M = (M_{m,n})$ equipped with two operations

$$\circ : M_{m,n} \times M_{n,k} \to M_{m,k}, \qquad m, n, k \geqslant 0$$

$$\Box : M_{m,n} \times M_{m',n'} \to M_{m+m',n+n'}, \qquad m, n, m', n' \geqslant 0$$

which are associative in the obvious way and satisfy the distributivity law

$$(f \circ g) \Box (f' \circ g') = (f \Box f') \circ (g \Box g')$$

whenever all the above operations are defined. Moreover, both the operations $\circ$ and $\Box$ are unitary, i.e., M is equipped with a sequence of constants $e_n \in M_{n,n}$ $(n \geqslant 0)$, called units, such that

$$e_m \circ f = f = f \circ e_n, \;\; e_0 \Box f = f = f \Box e_0$$

for all $f \in M_{m,n}$ and all $m, n \geqslant 0$, and the additional condition

$$e_m \Box e_n = e_{m+n}, \qquad \text{for all } m, n \geqslant 0$$

holds.

In other words a magmoid is nothing but an *x-category* (cf. [5, 8, 9]) or strict monoidal category (cf. Chapter VII of [11]) whose set of objects is the set of natural numbers. Submagmoids, morphisms, congruences and quotients of magmoids are defined in the obvious way.

Example 1 (Magmoids of Functions and Relations). *Let Q be a non-empty set and denote by $V_m(Q)$ the set of all vertical words of Q with length $m \geqslant 0$. Concatenation on vertical words is symbolized by $\Box$:*

$$\textit{if} \quad u = \begin{matrix} q_1 \\ \vdots \\ q_m \end{matrix} \quad \textit{and} \quad w = \begin{matrix} p_1 \\ \vdots \\ p_n \end{matrix}, \quad \textit{then} \quad u \Box w = \begin{matrix} q_1 \\ \vdots \\ q_m \\ p_1 \\ \vdots \\ p_n \end{matrix}$$

thus $\begin{matrix} q_1 \\ \vdots \\ q_m \end{matrix} = q_1 \Box \cdots \Box q_m$.

In what follows, for the sake of simplicity, we shall identify the vertical word $q_1 \Box \cdots \Box q_m$ with the word $q_1 \ldots q_m$.

The sets $Funct_{m,n}(Q)$ of all functions from $V_m(Q)$ to $V_n(Q)$

$$Funct_{m,n}(Q) = \{f \mid f : V_m(Q) \to V_n(Q)\}, \qquad m, n \geqslant 0,$$

can be structured into a magmoid with $\circ$ being the usual function composition, while the operation $\Box$ is the function boxing defined as follows: for $f \in Funct_{m,n}(Q)$ and $f' \in Funct_{m',n'}(Q)$

$$f \Box f'(u \Box u') = f(u) \Box f'(u'), \qquad u \in V_m(Q)\,,\, u' \in V_{m'}(Q).$$

In a similar way the sets

$$Rel_{m,n}(Q) = \{R \mid R \subseteq V_m(Q) \times V_n(Q)\}$$

of all relations from $V_m(Q)$ to $V_n(Q)$ can be organized into a magmoid, $\circ$ being the relation composition and $\square$ the relation concatenation.

Clearly $Funct(Q) = (Funct_{m,n}(Q))$ is a sub-magmoid of $Rel(Q) = (Rel_{m,n}(Q))$.

Example 2 (The Magmoid of Hypergraphs). *Given a finite alphabet X, we denote by X^* the set of all words over X and for every word $w \in X^*$, $|w|$ denotes its length. Formally, a* concrete (m, n)-graph *over a doubly ranked alphabet $\Sigma = (\Sigma_{m,n})$ is a tuple*

$$G = (V, E, s, t, l, begin, end)$$

where

- *V is the finite set of nodes,*
- *E is the finite set of hyperedges,*
- *$s : E \to V^*$ is the source function,*
- *$t : E \to V^*$ is the target function,*
- *$l : E \to \Sigma$ is the labelling function such that $rank(l(e)) = (|s(e)|, |t(e)|)$ for every $e \in E$,*
- *$begin \in V^*$ with $|begin| = m$ is the sequence of begin nodes and*
- *$end \in V^*$ with $|end| = n$ is the sequence of end nodes.*

For an edge e of a hypergraph G we simply write $rank(e)$ to denote $rank(l(e))$. We denote by $GR_{m,n}(\Sigma)$ the set of all $(m, n) - graphs$ labelled over Σ.

The specific sets V and E chosen to define a concrete graph G are actually irrelevant. We shall not distinguish between two isomorphic graphs. Hence we have the following definition of an abstract graph. Two concrete (m, n)-graphs $G = (V, E, s, t, l, begin, end)$ and $G' = (V', E', s', t', l', begin', end')$ over Σ are isomorphic iff there exist two bijections $h_V : V \to V'$ and $h_E : E \to E'$ commuting with source, target, labelling, begin and end in the usual way.

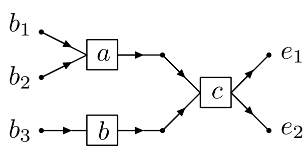
An abstract (m, n)-graph *is defined to be the equivalence class of a concrete (m, n)-graph with respect to isomorphism. We denote by $GR_{m,n}(\Sigma)$ the set of all abstract (m, n)-graphs over Σ. Since we shall mainly be interested in abstract graphs we shall simply call them graphs except when it is necessary to emphasize that they are defined up to an isomorphism.*

Any graph $G \in GR_{m,n}(\Sigma)$ having no edges, is called a discrete (m, n)-graph. *Given an edge label $\sigma \in \Sigma_{m,n}$, we still denote by σ the (m, n)-graph such that $V = \{x_1, \ldots, x_m, y_1, \ldots, y_n\}$, $E = \{e\}$ with $l(e) = \sigma$, $begin = s(e) = x_1 \cdots x_m$ and $end = t(e) = y_1 \cdots y_n$.*

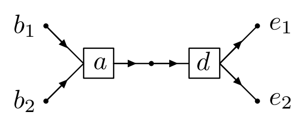
If G is a (m, n)-graph represented by $(V, E, s, t, l, begin, end)$ and H is an (n, k)-graph represented by $(V', E', s', t', l', begin', end')$ then their product $G \circ H$ *is the (m, k)-graph represented by the concrete graph obtained by taking the disjoint union of G and H and then identifying the ith end node of G with the ith begin node of H, for every $i \in \{1, ..., n\}$; also, $begin(G \circ H) = begin(G)$ and $end(G \circ H) = end(H)$.*

The sum $G \,\Box\, H$ *of arbitrary graphs G and H is their disjoint union with their sequences of begin nodes concatenated and similarly for their end nodes.*

For instance let $\Sigma = \{a, b, c, d\}$, with $rank(a) = (2, 1)$, $rank(b) = (1, 1)$, $rank(c) = (2, 2)$ and $rank(d) = (1, 2)$. In the following pictures, edges are represented by boxes, nodes by dots, and the sources and targets of an edge by directed lines that enter and leave the corresponding box, respectively. The order of the sources and targets of an edge is the vertical order of the directed lines as drawn in the pictures. We display two graphs $G \in GR_{3,2}(\Sigma)$ and $H \in GR_{2,2}(\Sigma)$, where the ith begin node is indicated by b_i, and the ith end node by e_i.

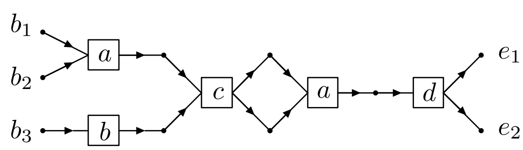


G

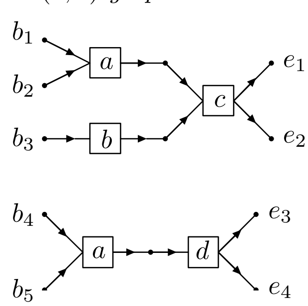


H

Then their product $G \circ H$ is the $(3, 2)$-graph



and, their sum $G \,\Box\, H$ is the $(5, 4)$-graph



For every $n \in \mathbb{N}$ the unit E_n of rank (n, n) is the discrete graph with nodes $x_1, ..., x_n$ and $begin(E_n) = end(E_n) = x_1 \cdots x_n$. Note that E_0 is the empty graph.

It is straightforward to verify that $GR(\Sigma) = (GR_{m,n}(\Sigma))$ with the operations defined above is a magmoid, see Lemma 6 of [6]. Subsets of $GR(\Sigma)$ are referred as graph languages. The discrete graphs of $GR(\Sigma)$ form manifestly a *sub-magmoid $DISC$ of $GR(\Sigma)$ and the function sending each graph $G \in GR(\Sigma)$ to its underlying discrete graph is indeed an epimorphism of magmoids*

$$disc_\Sigma : GR(\Sigma) \to DISC.$$

Let G be an $(m,n)-$ graph over a doubly ranked alphabet Σ and e an edge of G whose source and target sequences are $v_1, ..., v_\kappa$ and $u_1, ..., u_\lambda$ respectively. The triple (v_i, e, u_j) is a simple directed path of G from v_i to u_j $(1 \leq i \leq \kappa,\ 1 \leq j \leq \lambda)$; its label will be a new symbol σ_{ij}, where σ is the label of e. We set

$$path(\Sigma) = \{\sigma_{ij} \mid \sigma \in \Sigma_{m,n}, 1 \leq i \leq m, 1 \leq j \leq n\ ,\ m,n \geqslant 0\}.$$

In this way, we obtain a canonical doubly ranked function

$$path : GR(\Sigma) \to GR(path(\Sigma))$$

sending each $(m,n)-$ graph G over Σ to the $(m,n)-$ graph $path(G)$ over $path(\Sigma)$ obtained by substituting every hyperedge e of G by all directed edges as defined above. It is easily seen that

$$path(G \circ G') = path(G) \circ path(G') \text{ and } path(G \,\square\, G') = path(G) \,\square\, path(G')$$

i.e., path is actually a morphism of magmoids. The graph $G \in GR_{m,n}(\Sigma)$ is said to be connected whenever $path(G)$ is connected in the ordinary sense.

In order to fix our notation we need the following definitions. Let $M = M_{m,n}$ be a magmoid. We say that a doubly ranked family $L = (L_{m,n})$ is a *subset* of M (notation $L \subseteq M$), whenever $L_{m,n} \subseteq M_{m,n}$ for all $m,n \in \mathbb{N}$. The boolean operations on subsets of M are defined in the obvious way. A family $f = (f_{m,n})$ with $f_{m,n} \in M_{m,n}$ for all $m,n \in \mathbb{N}$ is called an *element* of M. A zero element in M is an element $0 = (0_{m,n})$ such that

$$0_{m,n} \circ M_{n,k} = 0_{m,k} = M_{m,n} \circ 0_{n,k} \quad \text{and}$$

$$0_{m,n} \,\square\, M_{p,q} = 0_{m+p,n+q} = M_{m,n} \,\square\, 0_{p,q}.$$

To any magmoid M we can adjoin a zero element in the obvious way. M^0 stands for the so obtained magmoid.

The structure of all subsets of a magmoid M is that of a double semiring with respect to the operations of union, $\circ$ *-product* and $\square$ *-product*. More precisely a *double semiring* is a 7-tuple $(K, +, 0, \circ, e, \square, f)$ where $\circ$ and $\square$ are two binary operations on K such that both $(K, +, 0, \circ, e)$ and $(K, +, 0, \square, f)$ are semirings with units e and f respectively and the following distributivity law is satisfied:

$$(a \circ b) \,\square\, (c \circ d) = (a \,\square\, c) \circ (b \,\square\, d), \qquad a,b,c,d \in K.$$

For instance, any commutative semiring can be viewed as a double semiring.

Next, given subsets L, L' of a magmoid M (with unit sequence e_n) we define their $\circ$ *-product* $L \circ L'$ by setting

$$(L \circ L')_{m,n} = \bigcup_{k \geqslant 0} L_{m,k} \circ L'_{k,n}, \qquad m,n \in \mathbb{N}$$

and their $\Box$ -*product* $L \Box L'$ by setting

$$(L \Box L')_{m,n} = \bigcup_{\substack{\kappa+\kappa'=m \\ \lambda+\lambda'=n}} L_{\kappa,\lambda} \Box L'_{\kappa',\lambda'}, \qquad m,n \in \mathbb{N}.$$

The subsets E and F of M given by $E_{m,n} = \{e_n\}$ if $m = n$ and $\emptyset$ else, while $F_{m,n} = \{e_0\}$ if $m = n = 0$ and $\emptyset$ else, are the units of the operations $\circ$ and $\Box$ respectively. The reader will verify that the set of all subsets of M together with $\cup, \circ, \Box$ is a double semiring.

The $\circ$ -*star* is the union of the successive $\circ$ -*powers* of $L \subseteq M$:

$$L^{\circ} = \bigcup_{k \geqslant 0} L^{\circ,k}$$

where $L^{\circ,k}$ is inductively given by

$$L^{\circ,0} = E, \;\; L^{\circ,1} = L, ..., L^{\circ,k+1} = L \circ L^{\circ,k}.$$

The $\Box$ -*star* $L^{\Box}$ is defined analogously.

3 Free magmoids

Let $X = (X_{m,n})$ be a doubly ranked alphabet. We denote by $\mathcal{F}(X) = (F_{m,n}(X))$ the set of expressions, built from the operations $\circ$ and $\Box$ and the constants in X, together with the units $e_n \in F_{n,n}(X)$, $n \geq 0$, where e_n is a specified symbol not belonging to $X_{n,n}$; we set $e_1 = e$. More precisely, $\mathcal{F}(X)$ is the smallest doubly ranked set inductively defined as follows:

1. $X_{m,n} \subseteq F_{m,n}(X)$, for all $m, n \geqslant 0$.
2. For every $m \geqslant 0$, $e_m \in F_{m,m}(X)$.
3. If $\alpha \in F_{m,n}(X)$ and $\beta \in F_{n,k}(X)$ then the expression $(\alpha \circ \beta)$ is in $F_{m,k}(X)$ for all $m, n, k \geqslant 0$.
4. If $\alpha \in F_{m,n}(X)$ and $\beta \in F_{m',n'}(X)$ then the expression $(\alpha \Box \beta)$ is in $F_{m+m',n+n'}(X)$ for all $m, n, m', n' \geqslant 0$.

Let $\sim = (\sim_{m,n})$ be the congruence on $\mathcal{F}(X)$ generated by the following relations :

$$f_1 \circ (f_2 \circ f_3) \sim (f_1 \circ f_2) \circ f_3 \text{ and } g_1 \Box (g_2 \Box g_3) \sim (g_1 \Box g_2) \Box g_3$$

$$(f_1 \circ f_2) \Box (g_1 \circ g_2) \sim (f_1 \Box g_1) \circ (f_2 \Box g_2)$$

$$e_m \circ f \sim f \sim f \circ e_n, \; e_0 \Box f \sim f \sim f \Box e_0, \; e_m \Box e_n \sim e_{m+n}$$

for all $f_1, f_2, f_3, g_1, g_2, g_3, f \in \mathcal{F}(X)$ of the appropriate rank and all $m, n \geq 0$.

The elements of the quotient $mag_{m,n}(X) = F_{m,n}(X)/\sim_{m,n}$ are called (m,n)-*patterns* over the alphabet X. Note that $mag(X) = (mag_{m,n}(X))$ is by construction a magmoid. Given a finite set $\{x_1, \ldots, x_n\}$ we denote by $mag(x_1, \ldots, x_n)$ the magmoid $mag(\{x_1, \ldots, x_n\})$.

Our patterns are exactly the unsorted abstract dags of [10] and [3]. For another formalization see also [7].

Remark. It was in [9], where graphs (viz. the derivation graphs of type-0 Chomsky grammars) were first characterized as a free x-category.

Convention. *From now on, the operations $\circ$ and $\square$ will also be denoted as horizontal and vertical concatenation:*

$$\alpha \circ \beta \text{ as } \alpha\beta\ , \text{ and } \alpha \square \beta \text{ as } \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

respectively. With this convention the distributivity law takes the form

$$\begin{pmatrix} \alpha_1\beta_1 \\ \alpha_2\beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

whereas the nth unit element e_n can be written

$$e_n = e \square e \square \cdots \square e = \begin{pmatrix} e \\ \vdots \\ e \end{pmatrix} \quad (n \text{ times}).$$

The construction of the free magmoid follows naturally.

Theorem 1. *Let $X = (X_{m,n})$ be a doubly ranked alphabet. The magmoid $mag(X)$ is the free magmoid generated by X. In other words, the injective function*

$$j : X \to mag(X), \quad j(x) = x, \text{ for all } x \in X,$$

has the following universal property: for any function $h : X \to M$, where M is a magmoid, there is a unique morphism of magmoids

$$\hat{h} : mag(X) \to M$$

rendering commutative the following diagram.

$$\begin{tikzcd} X \arrow[r, "j"] \arrow[dr, "h"'] & mag(X) \arrow[d, "\hat{h}"] \\ & M \end{tikzcd}$$

Explicitly $\hat{h}$ is inductively defined by

- $\hat{h}(x) = h(x)$, $x \in X$,
- $\hat{h}(p_1 p_2) = \hat{h}(p_1) \circ \hat{h}(p_2)$, $p_1 \in mag_{m,n}(X)$, $p_2 \in mag_{n,k}(X)$, $m, n, k \geq 0$,
- $\hat{h}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \hat{h}(p_1) \Box \hat{h}(p_2)$, $p_1 \in mag_{m,n}(X)$, $p_2 \in mag_{m',n'}(X)$, $m, n, m', n' \geq 0$.

Subsets of $mag(X)$ are called *pattern languages*. A pattern equation over the doubly ranked alphabet X is a pair (p, p') with $p, p' \in mag_{m,n}(X)$ for some $m, n \geq 0$. We say that the equation $p = p'$ is satisfied in the magmoid M, whenever for any function $h : X \to M$, we have $\hat{h}(p) = \hat{h}(p')$, where $\hat{h} : mag(X) \to M$ is the unique extension of h as described in Theorem 1.

Example 3. *Let M be a magmoid. The* valuation function

$$val_M : mag(M) \to M$$

is the morphism of magmoids uniquely extending the identity function

$$i : M \to M.$$

If α is a pattern constructed by elements of M, $val_M(\alpha)$ is obtained by performing in M all the indicated operations on it. For instance if $p \in M_{2,2}$ and $\omega \in M_{2,1}$ then

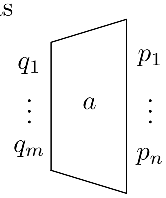
$$val \begin{pmatrix} p & p \\ p & \omega \end{pmatrix} = (p \circ p) \Box (p \circ \omega).$$

4 Pattern automata

Let $X = (X_{m,n})$ be a doubly ranked alphabet. A nondeterministic pattern automaton over X is a structure $\mathcal{A} = (X, Q, I, E, T)$, where Q is a finite set of states, $I \in \mathcal{P}(Q)$ and $T \in \mathcal{P}(Q)$ are the sets of initial and final states, and E is a finite set

$$E \subseteq \bigcup_{m,n \geq 0} V_m(Q) \times X_{m,n} \times V_n(Q).$$

The elements of E are called transitions ($V_m(Q)$ is the set of all vertical words of Q whose length is m, see Example 1). The triple $(u, a, v) \in E$ can be depicted as



$$u = q_1 \ldots q_m, \quad v = p_1, \ldots, p_n$$

In the case that for any two different triples (u, α, v) and (u', α, v') of E it holds $u \neq u'$ the automaton is called deterministic.

The behavior of the automaton is obtained as follows: we first define the doubly ranked function

$$ext_{m,n} : mag_{m,n}(E) \to V_m(Q) \times V_n(Q), \qquad m, n \geq 0,$$

by sending each element of $mag_{m,n}(E)$ to the begin and end state words of it:

- $ext_{m,n}(r) = (u, v)$ if $r = (u, a, v) \in E$, $a \in X_{m,n}$
- if $ext_{m,n}(s) = (u, v)$ and $ext_{n,k}(s') = (u', v')$, then $ext_{m,k}(s\,s') = (u, v')$
- if $ext_{m,n}(t) = (u, v)$ and $ext_{m',n'}(t') = (u', v')$, then $ext_{m+m',n+n'}\begin{pmatrix} t \\ t' \end{pmatrix} = (u\Box u', v\Box v')$, where $u\Box u'$ is the vertical concatenation of u and u'.

On the other hand we define the subset $L(\mathcal{A})$ of $mag(E)$ inductively by setting

- $r \in L_{m,n}(\mathcal{A})$ for all $r = (u, a, v) \in E, \quad a \in X_{m,n}$,
- if $s \in L_{m,n}(\mathcal{A})$, $s' \in L_{n,k}(\mathcal{A})$ and $ext_{m,n}(s) = (u, v)$, $ext_{n,k}(s') = (v, w)$ then $s\,s' \in L_{m,k}(\mathcal{A})$,
- If $t \in L_{m,n}(\mathcal{A})$, $t' \in L_{m',n'}(\mathcal{A})$ then $\begin{pmatrix} t \\ t' \end{pmatrix} \in L_{m+m',n+n'}(\mathcal{A})$.

Then we set

$$|\mathcal{A}| = \{s \in L(\mathcal{A}) \mid ext(s) \in I \times T\}.$$

Theorem 2. *Given a nondeterministic pattern automaton $\mathcal{A} = (X, Q, I, E, T)$ we can construct a deterministic automaton $\mathcal{A}'$ with the same behavior.*

Proof. We take the powerset $\mathcal{P}(Q)$ as the state set of $\mathcal{A}'$ and we define the set E_{det} as follows: for each $\alpha \in X_{m,n}$, the triple $(Q_1 \cdots Q_m, \alpha, P_1 \cdots P_n) \in E_{det}$ whenever there exist $q_1 \in Q_1, \ldots, q_m \in Q_m$ and $p_1 \in P_1, \ldots, p_n \in P_n$ such that $(q_1 \cdots q_m, \alpha, p_1 \cdots p_n) \in E$.

Moreover, we define the rational set $I' \subseteq \mathcal{P}(Q)^*$ by taking the word $Q_1 \cdots Q_m$ of $\mathcal{P}(Q)^*$ to be in I' whenever there exist $q_1 \in Q_1, \ldots, q_m \in Q_m$ such that the word $q_1 \cdots q_m \in I$. The set T' is obtained in a similar way. The sets I' and T' are rational subsets of $\mathcal{P}(Q)^*$ as confirms the lemma below, and we can verify that the so obtained deterministic automaton $\mathcal{A}'$ is equivalent to $\mathcal{A}$.

$\square$

Lemma 1. *If $L \subseteq A^*$ is recognizable, then so is $\hat{L} \subseteq \mathcal{P}(A)^*$ with $Q_1 \cdots Q_m \in \hat{L}$ whenever there are $q_1 \in Q_1, \ldots, q_m \in Q_m$ such that $q_1 \cdots q_m \in L$.*

Proof. Let $h : A^* \to M$ be a monoid morphism (M finite monoid) so that $L = h^{-1}(P)$ for some $P \subseteq M$. Let us consider the powerset monoid $\mathcal{P}(M)$ whose operation is the subset multiplication and the monoid morphism $\hat{h} : \mathcal{P}(A)^* \to \mathcal{P}(M)$ defined by the formula

$$\hat{h}(Q_1 \cdots Q_m) = h(Q_1) \bullet \cdots \bullet h(Q_m)$$

where "$\bullet$" is the multiplication of $\mathcal{P}(M)$. Then $\hat{L} = \hat{h}^{-1}(P')$, where

$$P' = \{B \mid B \subseteq M, B \cap P \neq 0\}.$$

$\square$

The set of all behaviors of (deterministic or nondeterministic) pattern automata over X is denoted $AREC(X)$.

Remark. Notice that our nondeterministic pattern automata where first introduced by Bossut, Dauchet and Warin in [3], under the name pdag (planar directed acyclic graph) automata. In [3] it is proved that $AREC(X)$ is closed under union, nondeterministic parallel composition, serial composition, and the iterations of these compositions.

Proposition 1. *If $F : mag(X) \to mag(X')$ is a magmoid morphism, then $L \in AREC(X)$ implies $F(L) \in AREC(X')$.*

Proof. Assume that $\mathcal{A} = (X, Q, I, E, T)$ is a pattern automaton over X; then the automaton

$$F(\mathcal{A}) = (X', Q, I, F(E), T)$$

with

$$F(E) = \{(u, F(a), v) \mid (u, a, v) \in E\}$$

computes the subset $F(|\mathcal{A}|)$. $\square$

Proposition 2. *$AREC(X)$ is closed under $\square$ and $\square$-star*

Proof. Let $\mathcal{A}^i = (X, Q^i, I^i, E^i, T^i)$, $i = 1, 2$ be two pattern automata, with $Q^1 \cap Q^2 = \emptyset$. Then $|\mathcal{A}^1|\square|\mathcal{A}^2|$ is the behavior of the automaton

$$\mathcal{A}^1 \square \mathcal{A}^2 = (X, Q^1 \cup Q^2, I^1 I^2, E^1 \cup E^2, T^1 T^2)$$

where $I^1 I^2$ (resp. $T^1 T^2$) is the concatenation of $I^1, I^2 \in \mathcal{P}(Q)$ (resp. $T^1, T^2 \in \mathcal{P}(Q)$). Moreover, for any automaton $\mathcal{A} = (X, Q, I, E, T)$ the automaton $\mathcal{A}^{\square} = (X, Q, I^*, E, T^*)$ recognizes $L(\mathcal{A})^{\square}$. $\square$

Proposition 3. *If $L_1, L_2 \in AREC(X)$ then $L_1 \cap L_2 \in AREC(X)$ furthermore if $H : mag(X) \to mag(Y)$ is an alphabetic morphism then $L \in AREC(Y)$ implies $H^{-1}(L) \in AREC(X)$.*

Proof. A product construction is used. Let $\mathcal{A}_i = (X, Q_, I_i, E_i, T_i)$ be a pattern automaton and denote by $pr_i : Q_1 \times Q_2 \to Q_i$ the canonical projection and $pr_i^* : (Q_1 \times Q_2)^* \to Q_i^*$ its extension to free monoids (i=1,2). The automaton $\mathcal{A}_1 \cap \mathcal{A}_2 = (X, Q_1 \times Q_2, I, E, T)$ where

$$I = (pr_1^*)^{-1}(I_1) \cap (pr_2^*)^{-1}(I_2)$$

and

$$T = (pr_1^*)^{-1}(T_1) \cap (pr_2^*)^{-1}(T_2)$$

whereas $(w,x,v) \in E$ if and only if

$$(pr_1^*(w), x, pr_1^* v) \in E_1 \text{ and } (pr_2^*(w), x, pr_2^* v) \in E_2$$

has as behavior $|\mathcal{A}_1 \cap \mathcal{A}_2| = |\mathcal{A}_1| \cap |\mathcal{A}_2|$.

Next consider an automaton $\mathcal{A}$ over the doubly ranked alphabet $Y = (Y_{m,n})$. Then the automaton $H^{-1}(\mathcal{A}) = (X, Q, I, \bar{E}, T)$ with

$$\bar{E} = \{(u,x,v) \mid (u, H(x), v) \in E, x \in X\}$$

computes the set $H^{-1}(|\mathcal{A}|)$. □

5 D-magmoids

As we have seen in the introduction, the equations $\mathcal{E}$ are satisfied in $GR(\Sigma)$ by replacing π by Π and $i_{\kappa\lambda}$ by $I_{\kappa,\lambda}$. Magmoids with such a property are called *D-magmoids*. More precisely, a D-magmoid $\mathcal{M} = (M, D)$ consists of a magmoid M, called the domain of the D-magmoid, and a set $D = \{d, d_{01}, d_{21}, d_{10}, d_{12}\}$, where $d \in M_{2,2}$, $d_{01} \in M_{0,1}$, $d_{21} \in M_{2,1}$, $d_{10} \in M_{1,0}$, $d_{12} \in M_{1,2}$ and the equations $\mathcal{E}$ are satisfied by replacing π with d, $i_{\kappa\lambda}$ by $d_{\kappa\lambda}$ and the letters $\sigma \in \Sigma_{m,n}$ by all elements $\alpha \in M_{m,n}$, $m,n \geq 0$, i.e., the following 15 equations hold:

$$\mathcal{E}_D: \quad d\,d = e_2, \quad \begin{pmatrix} e \\ d \end{pmatrix} \begin{pmatrix} d \\ e \end{pmatrix} \begin{pmatrix} e \\ d \end{pmatrix} = \begin{pmatrix} d \\ e \end{pmatrix} \begin{pmatrix} e \\ d \end{pmatrix} \begin{pmatrix} d \\ e \end{pmatrix}, \quad \begin{pmatrix} e \\ d_{21} \end{pmatrix} d_{21} = \begin{pmatrix} d_{21} \\ e \end{pmatrix} d_{21},$$

$$\begin{pmatrix} e \\ d_{01} \end{pmatrix} d_{21} = e, \quad d\,d_{21} = d_{21}, \quad \begin{pmatrix} d \\ e \end{pmatrix} \begin{pmatrix} e \\ d \end{pmatrix} \begin{pmatrix} d_{21} \\ e \end{pmatrix} = \begin{pmatrix} e \\ d_{21} \end{pmatrix} d, \quad \begin{pmatrix} e \\ d_{01} \end{pmatrix} d = \begin{pmatrix} d_{01} \\ e \end{pmatrix},$$

$$d_{12} \begin{pmatrix} e \\ d_{12} \end{pmatrix} = d_{12} \begin{pmatrix} d_{12} \\ e \end{pmatrix}, \quad d_{12} \begin{pmatrix} e \\ d_{10} \end{pmatrix} = e, \quad d_{12} d = d_{12}, \quad d_{12}\, d_{21} = e,$$

$$\begin{pmatrix} d_{12} \\ e \end{pmatrix} \begin{pmatrix} e \\ d \end{pmatrix} \begin{pmatrix} d \\ e \end{pmatrix} = d \begin{pmatrix} e \\ d_{12} \end{pmatrix}, \quad d \begin{pmatrix} e \\ d_{10} \end{pmatrix} = \begin{pmatrix} d_{10} \\ e \end{pmatrix}, \quad \begin{pmatrix} d_{12} \\ e \end{pmatrix} \begin{pmatrix} e \\ d_{21} \end{pmatrix} = d_{21}\, d_{12},$$

$$d_{m,1} \begin{pmatrix} \alpha \\ e \end{pmatrix} = \begin{pmatrix} e \\ \alpha \end{pmatrix} d_{n,1}, \quad \text{for all } \alpha \in M_{m,n}, \ \ m,n \geq 0,$$

where $d_{m,1}$ is defined analogously with $\pi_{m,1}$.

We recall from [4] that, if the last equation holds in $GR(\Sigma)$ for all the letters of the doubly ranked alphabet Σ, then it holds for every element of $GR(\Sigma)$. Thus the pair $(GR(\Sigma), D)$, where $D = \{\Pi, I_{0,1}, I_{2,1}, I_{1,0}, I_{1,2}\}$ is a D-magmoid.

A D-magmoid is called relational whenever its domain is the magmoid of relations $Rel(Q)$, where Q is an arbitrary set.

Given D-magmoids (M, D) and (M', D'), a magmoid morphism $H : M \to M'$ such that $H(d) = d'$ and $H(d_{\kappa\lambda}) = d'_{\kappa\lambda}$, is called a morphism of D-magmoids.

Next we present two relational D-magmoids which will be useful in the construction of graph automata.

Diagonal D-magmoids. Given a finite set Q, we consider the following elements of $Rel(Q)$:

- $d \in Rel_{2,2}(Q)$, with $d = \{(q_1q_2, q_2q_1) \mid q_1, q_2 \in Q\}$,
- $d_{01} \in Rel_{0,1}(Q)$, with $d = \{(\varepsilon, q) \mid q \in Q\}$,
- $d_{21} \in Rel_{2,1}(Q)$, with $d = \{(qq, q) \mid q \in Q\}$,
- $d_{10} \in Rel_{1,0}(Q)$, with $d = \{(q, \varepsilon) \mid q \in Q\}$,
- $d_{12} \in Rel_{1,2}(Q)$, with $d = \{(q, qq) \mid q \in Q\}$.

The pair $\mathcal{R}_\Delta = (Rel(Q), \{d, d_{01}, d_{21}, d_{10}, d_{12}\})$ manifestly constitutes a D-magmoid.

Group D-magmoids. Let $G = (G, +, 0)$ be an additive group, we define the following elements of $Rel(G)$:

- $d \in Rel_{2,2}(G)$, with $d = \{(g_1g_2, g_2g_1) \mid g_1, g_2 \in G\}$,
- $d_{01} \in Rel_{0,1}(G)$, with $d = \{(\varepsilon, 0)\}$, where ε is the empty word,
- $d_{21} \in Rel_{2,1}(G)$, with $d = \{(g_1g_2, g) \mid g_1 + g_2 = g \text{ and } g, g_1, g_2 \in G\}$,
- $d_{10} \in Rel_{1,0}(G)$, with $d = \{(0, \varepsilon)\}$, where ε is the empty word,
- $d_{12} \in Rel_{1,2}(G)$, with $d = \{(g, g_1g_2) \mid g = g_1 + g_2 \text{ and } g, g_1, g_2 \in G\}$.

It is not hard to verify, that the equations $\mathcal{E}_D$ are satisfied inside the magmoid $Rel(G)$. Thus the pair $\mathcal{R}_G = (Rel(G), D)$, with $D = \{d, d_{01}, d_{21}, d_{10}, d_{12}\}$, forms manifestly a D-magmoid.

We have already discussed how the set $GR(\Sigma)$ can be structured into a D-magmoid; in fact it is the free D-magmoid generated by Σ.

Theorem 3. *The doubly ranked function $J : \Sigma \to GR(\Sigma)$, with $J(\sigma) = \sigma$, for all $\sigma \in \Sigma$, has the following universal property: for any doubly ranked function $f : \Sigma \to \mathcal{M}$, where $\mathcal{M}$ is a D-magmoid, there exists a unique morphism of D-magmoids $\bar{f} : GR(\Sigma) \to \mathcal{M}$ making commutative the triangle*

$$\begin{tikzcd} \Sigma \arrow[r, "J"] \arrow[dr, "f"'] & GR(\Sigma) \arrow[d, "\bar{f}"] \\ & \mathcal{M} \end{tikzcd}$$

Proof. By virtue of Theorem 1 there exists a unique magmoid morphism $\hat{f}$ making commutative the triangle:

$$\begin{tikzcd}
\Sigma \arrow[r, "j"] \arrow[dr, "f"'] & mag(\Sigma \cup D) \arrow[d, "\hat{f}"] \\
& M
\end{tikzcd}$$

Since all equations $\mathcal{E}$ are valid in $\mathcal{M}$, the kernel of $\hat{f}$ includes $\mathcal{E}$: $Ker(\hat{f}) \supseteq \mathcal{E}$. It turns out that $\hat{f}$ induces a unique D-magmoid morphism

$$\bar{f} : mag(\Sigma \cup D)/\mathcal{E} = GR(\Sigma) \longrightarrow \mathcal{M}$$

rendering commutative the triangle:

$$\begin{tikzcd}
mag(\Sigma \cup D) \arrow[r, "g"] \arrow[dr, "\hat{f}"'] & GR(\Sigma) \arrow[d, "\bar{f}"] \\
& \mathcal{M}
\end{tikzcd}$$

The result comes by combining the above two diagrams. □

6 Graph Automata

The algebraic structure of D-magmoids, and in particular relational D-magmoids, will allow us to define automata on graphs. We will investigate the properties of two distinguished graph language recognition modes according to the D-magmoid structure that will be used as the state set. Hence we will speak of the diagonal and the group recognition mode of graph languages.

Let Σ be a doubly ranked alphabet. From now on we denote by A either a finite set Q or a finite group G. A graph automaton over the relational D-magmoid $\mathcal{R} = (Rel(A), D)$ is a structure $\mathcal{A} = (\mathcal{R}, A, I, E, T)$ where A is a finite set of states, $I, T \in Rat(A^*)$ are the initial and final rational state word languages and E is a finite set of transitions

$$E \subseteq \bigcup_{m,n \geq 0} V_m(A) \times \Sigma_{m,n} \times V_n(A).$$

The behavior of such an automaton is obtained as follows: the doubly ranked transition function of $\mathcal{A}$

$$\delta_{\mathcal{A}} : \Sigma \to Rel(A), \quad \delta_{\mathcal{A}}(\sigma) = \{(u, v) \mid (u, \sigma, v) \in E\}$$

is uniquely extended by Theorem 3 to a morphism of D-magmoids

$$\bar{\delta}_{\mathcal{A}} : GR(\Sigma) \to \mathcal{R}.$$

We set $|\mathcal{A}| = \bar{\delta}_{\mathcal{A}}^{-1}(F)$, where $F = (F_{m,n})$ is the doubly ranked set given by

$$F_{m,n} = (I \cap V_m(A)) \times (T \cap V_n(A)).$$

A graph language is called *recognizable* whenever it is obtained as the behavior of a graph automaton. The class of all recognizable graph languages over the doubly ranked alphabet Σ is denoted $Rec(\Sigma)$.

The above definition clearly shows that the selection of the D-magmoid structure plays an important role in the recognition procedure of the graph automaton.

The first mode of recognizability we introduce is the *diagonal recognizability*, obtained by virtue of the diagonal D-magmoid $\mathcal{R}_\Delta$. The class of all diagonal recognizable graph languages, over the doubly ranked alphabet Σ, is denoted Δ-$Rec(\Sigma)$.

Example 4. *We recall from the introduction that $I_{p,q}$ denotes the discrete (p, q)-graph having a single node x and whose begin and end sequences are $x \cdots x$ (p times) and $x \cdots x$ (q times) respectively. Notice that for all $n \geq 2$ it holds*

$$I_{1,n} = I_{1,2} \begin{pmatrix} I_{1,n-1} \\ E \end{pmatrix} = I_{1,2} \begin{pmatrix} E \\ I_{1,n-1} \end{pmatrix}$$

and similarly for $I_{n,1}$. Let $\sigma \in \Sigma_{1,1}$, the graph automaton

$$\mathcal{A}_p = (\mathcal{R}_\Delta, Q, I, E, T)$$

with $Q = \{q_1, q_2\}$, $I = \{q_1\}$, $T = \{q_2\}$ and $E = \{(q_1, \sigma, q_2)\}$, clearly computes the graph language L_p, consisting of the graphs:

$$\sigma, \quad I_{1,2} \begin{pmatrix} \sigma \\ \sigma \end{pmatrix} I_{2,1}, \quad I_{1,3} \begin{pmatrix} \sigma \\ \sigma \\ \sigma \end{pmatrix} I_{3,1}, \ \ldots$$

Let G be a finite, commutative group, we say that a graph language is G-*recognizable* whenever it is obtained as the behavior of a graph automaton over the group D-magmoid $\mathcal{R}_G$. The class of all G-recognizable graph languages, over the doubly ranked alphabet Σ, is denoted G-$Rec(\Sigma)$.

In the present paper we will mainly focus our attention in the case of $\mathcal{Z}_m$-recognizable graph languages, where $\mathcal{Z}_m = (\{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}, +, \overline{0})$ is the cyclic group of *modm* integers.

Example 5. *Let $m \geq 2$ and $\sigma \in \Sigma_{1,1}$, we denote by $L_{\sigma,m}$ the graph language consisting of all $(1, 1)$-graphs G in which the number of occurrences of σ in their edges (denoted $|G|_\sigma$) is a multiple of m, i.e.,*

$$L_{\sigma,m} = \{G \mid G \in GR(\Sigma),\ |G|_\sigma = m \cdot k,\ k \geq 1\}.$$

We construct the automaton $\mathcal{A}_m = (\mathcal{R}_{\mathcal{Z}_m}, \mathcal{Z}_m, I, E, T)$, where $I = \overline{0}^*$, $T = \overline{0}^*$ and $E = \{(k, \sigma, k+1) \mid \overline{0} \leq k \leq \overline{m-1}\}$. It is not hard to see that $|\mathcal{A}_m| = L_{\sigma,m}$ and thus for all $m \geq 2$, the language $L_{\sigma,m}$ is $\mathcal{Z}_m$-recognizable.

7 Closure properties of recognizable graph languages

The following propositions state that the two modes of graph recognizability we have introduced, are closed under union and intersection.

Proposition 4. *If $L_1, L_2 \in \Delta$-$Rec(\Sigma)$, then $L_1 \cap L_2, L_1 \cup L_2 \in \Delta$-$Rec(\Sigma)$.*

Proof. Let $\mathcal{A}_i = (\mathcal{R}_\Delta, Q, I_i, E_i, T_i)$, $i = 1, 2$, be two graph automata with behaviors $|\mathcal{A}_i| = L_i$, we set

$E' = \{((u_1, u_2), \sigma, (v_1, v_2)) \mid (u_i, \sigma, v_i) \in E_i, \sigma \in \Sigma_{m,n}, u_i \in Q^m, v_i \in Q^n, i = 1, 2\}.$

It is straightforward to verify that the structures

$$\mathcal{A}_1 \cap \mathcal{A}_2 = (\mathcal{R}_\Delta, Q \times Q, I_1 \times I_2, E', T_1 \times T_2)$$

and

$$\mathcal{A}_1 \cap \mathcal{A}_2 = (\mathcal{R}_\Delta, Q \times Q, I_1 \times I_2, E', T_1 \times Q^* \cup Q^* \times T_2)$$

constitute two graph automata with behaviors

$$|\mathcal{A}_1 \cap \mathcal{A}_2| = L_1 \cap L_2 \quad \text{and} \quad |\mathcal{A}_1 \cup \mathcal{A}_2| = L_1 \cup L_2.$$

$\square$

Proposition 5. *Let G and H be two finite commutative groups, if $L_1 \in G$-$Rec(\Sigma)$ and $L_2 \in H$-$Rec(\Sigma)$, then $L_1 \cup L_2, L_1 \cap L_2 \in (G \times H)$-$Rec(\Sigma)$. In particular, if $L_1, L_2 \in G$-$Rec(\Sigma)$ then $L_1 \cup L_2, L_1 \cap L_2 \in G$-$Rec(\Sigma)$.*

Proof. Similar with the proof of Proposition 4. $\square$

Proposition 6. *Consider the canonical magmoid morphism*

$$val_{\Sigma \cup D} : \ mag(\Sigma \cup D) \longrightarrow GR(\Sigma).$$

If $L \subseteq GR(\Sigma)$ is a recognizable language, then $val^{-1}_{\Sigma \cup D}(L)$ is a recognizable pattern language.

Proof. Let $\mathcal{A} = (\mathcal{R}, A, I, E, T)$ be a graph automaton where A is either a finite commutative group G or a finite set Q. By extending the transition function $\delta_\mathcal{A} : \Sigma \to Rel(A)$, we get a D-magmoid morphism

$$\bar{\delta}_\mathcal{A} : GR(\Sigma) \to \mathcal{R}.$$

Composing $\bar{\delta}_\mathcal{A}$ with $val_{\Sigma \cup D}$, we get the magmoid morphism

$$mag(\Sigma \cup D) \xrightarrow{val_{\Sigma \cup D}} GR(\Sigma) \xrightarrow{\bar{\delta}_\mathcal{A}} \mathcal{R}. \tag{1}$$

Therefore, the pattern automaton $\mathcal{B} = (\Sigma, A, I, B, T)$, where B is the restriction of (1) on $\Sigma \cup D$, recognizes $val^{-1}_{\Sigma \cup D}(L)$, as wanted. $\square$

Consider two doubly ranked alphabets Σ and Γ. By Theorem 3 any doubly ranked function $h : \Sigma \to GR(\Gamma)$ can be uniquely extended into a morphism of D-magmoids

$$\bar{h} : GR(\Sigma) \longrightarrow GR(\Gamma).$$

Such morphisms are called *graph homomorphisms*. In the case that $h(\Sigma) \subseteq \Gamma$, $\bar{h}$ is called *alphabetic*.

Using a technique similar with that of the previous proposition we get the following proposition.

Proposition 7. *Let*

$$\bar{h} : GR(\Sigma) \to GR(\Gamma)$$

be a graph homomorphism, if the language $L \subseteq GR(\Gamma)$ *is recognizable, then* $\bar{h}^{-1}(L) \subseteq GR(\Sigma)$ *is recognizable as well.*

8 Comparison of recognizability modes

Our first task is to determine a hierarchy inside the class of group recognizable graph languages.

Theorem 4. *If G and H are two finite groups and $\phi : G \to H$ a group epimorphism, then H-$Rec(\Sigma) \subseteq G$-$Rec(\Sigma)$.*

Proof. First we observe that the group epimorphism ϕ respects the elements $D = \{d, d_{01}, d_{21}, d_{10}, d_{12}\}$ and $D' = \{d', d'_{01}, d'_{21}, d'_{10}, d'_{12}\}$ of the D-magmoids $\mathcal{R}_G = (Rel(G), D)$ and $\mathcal{R}_H = (Rel(H), D')$, i.e., for instance we have

$$(q, q_1 q_2) \in d_{12} \quad \text{whenever} \quad (\phi(q), \phi(q_1)\phi(q_2)) \in d'_{12}$$

and similarly for d, d_{01}, d_{21} and d_{10}.

Now given a graph language $L \in H$-$Rec(\Sigma)$, let $\mathcal{A} = (\mathcal{R}_H, H, I, E, T)$ be the automaton recognizing L. We set $I' = \phi^{-1}(I)$, $T' = \phi^{-1}(T)$ and

$$E' = \{(s_1 \cdots s_m, \sigma, r_1 \cdots r_n) \mid (\phi(s_1) \cdots \phi(s_m), \sigma, \phi(r_1) \cdots \phi(r_n)) \in E, \;\; \sigma \in \Sigma\}.$$

Then the graph automaton $\mathcal{A}' = (\mathcal{R}_G, G, I', E', T')$ computes the language L. $\square$

Corollary 1. *Let G and H be two isomorphic finite groups, then G-$Rec(\Sigma) = H$-$Rec(\Sigma)$.*

Corollary 2. *Let $S = (\{0\}, +, 0)$ be the trivial group, then for all finite commutative groups G, it holds: S-$Rec(\Sigma) \subseteq G$-$Rec(\Sigma)$.*

Remark. The set $GR(\Sigma)$ of all graphs over the doubly ranked alphabet Σ belongs to the class S-$Rec(\Sigma)$. Indeed let $\mathcal{A} = (\mathcal{R}_S, S, I, E, T)$ be the graph automaton over the group D-magmoid $\mathcal{R}_S$ and the doubly ranked alphabet Σ, with $I = T = 0^*$ and $(0^m, \sigma, 0^n) \in E$ for all $\sigma \in \Sigma_{m,n}$ and $m, n \geq 0$. Clearly $|\mathcal{A}| = GR(\Sigma)$.

Corollary 3. *For all finite commutative groups G and H it holds*

$$G\text{-}Rec(\Sigma) \cup H\text{-}Rec(\Sigma) \subseteq (G \times H)\text{-}Rec(\Sigma).$$

Proof. By applying the previous theorem to the epimorphisms

$$pr_G : G \times H \to G, \quad pr_G(g,h) = g \quad \text{and} \quad pr_H : G \times H \to H, \quad pr_H(g,h) = h$$

we get G-$Rec(\Sigma) \subseteq (G \times H)$-$Rec(\Sigma)$ and H-$Rec(\Sigma) \subseteq (G \times H)$-$Rec(\Sigma)$, hence

$$G\text{-}Rec(\Sigma) \cup H\text{-}Rec(\Sigma) \subseteq (G \times H)\text{-}Rec(\Sigma).$$

□

Now taking into account that every finite commutative group G can be written as a direct sum $G = \mathcal{Z}_{m_1} \oplus \cdots \oplus \mathcal{Z}_{m_k}$, where $m_1, \ldots m_k \geq 2$, we get the following corollary.

Corollary 4. *For every commutative group G there exist $m_1, \ldots, m_k \geq 2$, such that*

$$\mathcal{Z}_{m_1}\text{-}Rec(\Sigma) \cup \cdots \cup \mathcal{Z}_{m_k}\text{-}Rec(\Sigma) \subseteq G\text{-}Rec(\Sigma).$$

Proposition 8. *Let $2 \leq m \leq n$, then $m \mid n$ if and only if $\mathcal{Z}_m$-$Rec(\Sigma) \subseteq \mathcal{Z}_n$-$Rec(\Sigma)$.*

Proof. First we assume that $m \mid n$, then $\mathcal{Z}_m$ is a subgroup of $\mathcal{Z}_n$ and there exists a group epimorphism $\phi : \mathcal{Z}_n \to \mathcal{Z}_m$ sending each element of $\mathcal{Z}_n$ to its modulo m class. Hence by Theorem 4, we obtain $\mathcal{Z}_m$-$Rec(\Sigma) \subseteq \mathcal{Z}_n$-$Rec(\Sigma)$.

Now let $m \nmid n$, and consider the graph language $L_{\sigma,m} \in GR_{1,1}(\Sigma)$ of Example 5. As we have seen, this graph language is $\mathcal{Z}_m$-recognizable. Assume that it is $\mathcal{Z}_n$-recognizable as well, and let $\mathcal{A} = (\mathcal{R}_{\mathcal{Z}_n}, \mathcal{Z}_n, I, E, T)$ be a graph automaton computing $L_{\sigma,m}$. If $(\overline{p}, \sigma, \overline{q}) \in E$ and $G \in |\mathcal{A}|$ is a graph with k σ's, where $m \mid k$, then the graph

$$H = I_{1,2} \begin{pmatrix} I_{1,n} \begin{pmatrix} \sigma \\ \vdots \\ \sigma \end{pmatrix} I_{n,1} \\ G \end{pmatrix} I_{2,1}$$

belongs to the behavior of $\mathcal{A}$, since $\mathcal{Z}_n$ is a group of order n, and thus $\overline{p}+\cdots+\overline{p} = n\cdot\overline{p} = \overline{0}$ and $\overline{q}+\cdots+\overline{q} = n\cdot\overline{q} = \overline{0}$. The graph H has $k+n$ σ's but $m \mid k$ and $m \nmid n$ and hence $m \nmid k + n$, which is a contradiction. Therefore, $\mathcal{Z}_m$-$Rec(\Sigma) \nsubseteq \mathcal{Z}_n$-$Rec(\Sigma)$ and this concludes the proof. □

The following proposition states that the two modes of graph recognizability we have introduced are incomparable.

Proposition 9. *For every $m \geq 2$, the classes $\mathcal{Z}_m$-$Rec(\Sigma)$ and Δ-$Rec(\Sigma)$ are incomparable.*

Proof. First we shall prove that, for every $m \geq 2$, the languages $L_{\sigma,m}$ of Example 5 are not diagonal recognizable. Assume that there exists a diagonal automaton $\mathcal{A}$ recognizing $L_{\sigma,m}$ and let G be a graph with m σ's. Then the graph G' that is obtained by replacing one edge of G, that is labelled by σ, with the $(1,1)$-graph

$$I_{1,2}\begin{pmatrix}\sigma\\ \sigma\end{pmatrix}I_{2,1}$$

should also belong to the behavior of $\mathcal{A}$. This is a contradiction since G' has $m+1$ edges labelled by σ.

Now consider the graph language L_p of Example 4, we shall prove that $L_p \notin \mathcal{Z}_2\text{-}Rec(\Sigma)$, the proof for $m \geq 2$ is analogous. Assume that $L_p \in \mathcal{Z}_2\text{-}Rec(\Sigma)$ and let $\mathcal{A} = (\mathcal{R}_{\mathcal{Z}_2}, \mathcal{Z}_2, I, E, T)$ be a graph automaton computing L_p. If the triple $(p, \sigma, p) \in E$, where $p = \overline{0}$ or $\overline{1}$, then the graph $\sigma\sigma$ should also belong to L_p, which is a contradiction. If moreover, $(\overline{0}, \sigma, \overline{0}) \notin E$, $(\overline{1}, \sigma, \overline{1}) \notin E$ and the triple $(\overline{0}, \sigma, \overline{1}) \in E$ (or $(\overline{1}, \sigma, \overline{0}) \in E$), then since the graph

$$I_{1,2}\begin{pmatrix}\sigma\\ \sigma\end{pmatrix}I_{2,1} \in L_p$$

we have $\overline{0} \in I$ and $\overline{0} \in T$, and hence the graph

$$I_{1,2}\begin{pmatrix}\sigma\\ \sigma\end{pmatrix}I_{2,1}I_{1,2}\begin{pmatrix}\sigma\\ \sigma\end{pmatrix}I_{2,1} \in |\mathcal{A}|$$

which is not the case. □

An immediate question, that arises from the previous proposition, is whether or not the classes Δ-*Rec*(Σ) and G-*Rec*(Σ) (where G is a finite, commutative group) are disjoint.

Remark. The set of all S-recognizable graph languages is equal with the set of all diagonal recognizable graph languages that are obtained as behaviors of automata with a unitary state set.

From the previous remark and Corollary 2 we deduce that, for all finite, commutative groups G, the intersection of Δ-*Rec*(Σ) and G-*Rec*(Σ), contains S-*Rec*(Σ) and hence the two classes are not disjoint.

References

[1] A. Arnold and M. Dauchet, *Théorie des magmoides. I.* RAIRO Inform. Théor. 12 (3), 235-257 (1978).

[2] A. Arnold and M. Dauchet, *Théorie des magmoides. II.* RAIRO Inform. Théor. 13 (2), 135-154 (1979).

[3] F. Bossut, M. Dauchet, B.Warin, *A Kleene theorem for a class of planar acyclic graphs.* Inform. and Comput. 117, 251-265 (1995).

[4] S.Bozapalidis and A.Kalampakas, *An Axiomatization of Graphs.* Acta Informatica 41, 19 - 61 (2004).

[5] V.Claus, *Ein Vollständigkeitssatz für Programme und Schaltkreise.* Acta Informatica 1, 64-78, (1971).

[6] J. Engelfriet and J.J. Vereijken, *Context-free graph grammars and concatenation of graphs.* Acta Informatica 34, 773-803 (1997).

[7] J.Gibbons, *An initial-algebra approach to directed acyclic graphs.* Mathematics of program construction (Kloster Irsee, 1995), 282-303, LNCS 947, Springer, Berlin, (1995).

[8] G.Hotz, *Eine Algebraisierung des Syntheseproblems von Schaltkreisen.* EIK 1, 185-205, 209-231, (1965).

[9] G.Hotz, *Eindeutigkeit und Mehrdeutigkeit formaler Sprachen.* EIK 2, 235-246, (1966).

[10] T. Kamimura, G. Slutzki, *Parallel and two-way automata on directed ordered acyclic graphs.* Inform. and Control 49, 10-51, (1981).

[11] S. MacLane, *Categories for the working mathematician.* Springer Verlag, (1971).

[12] D. Perrin, (Chapter 1) *Finite Automata*, Handbook of Theoretical Computer Science, Vol. B, J. van Leeuwen (ed.), MIT Press, (1990).