Algebra and Discrete Mathematics RESEARCH ARTICLE
Volume 11 **(2011)**. Number 2. pp. 51 – 63

Partial actions and automata

M. Dokuchaev, B. Novikov, G. Zholtkevych

Abstract. We use the notion of a partial action of a monoid to introduce a generalization of automata, which we call "a preautomaton". We study properties of preautomata and of languages recognized by preautomata.

Introduction

The concept of the partial action has been introduced in [2] for groups and extended to monoids in [7]. Therefore it is natural to investigate the influence in Automata Theory of the replacement of the full action of a free monoid by a partial action. Our article is devoted to the study of this topic.

We propose the term "preaction" for a notion which is called "partial action" in [7] and "strong partial action" in [4]. A generalization of the notion of an automaton which appears here, will be called "a preautomaton". This change of the terminology is caused by the fact that the term "partial automaton" is widely used in Automata Theory in a different sense (see, e.g., [3]).

We shall use the following facts from Automata Theory:

the Kleene theorem on regular languages,
the Myhill–Nerode theorem on languages and congruences,
the Eilenberg theorem on prefix decomposition.

All of them can be found in [1, 3, 5] etc.

This paper was partially supported by CNPq and FAPESP (Brazil).
2000 Mathematics Subject Classification: *20M30, 20M35, 68Q70.*

Key words and phrases: *Partial action, monoid, automaton, recognizable language, prefix code.*

The paper begins with a preliminary section on preactions. Then we define a preautomaton (Section 2) and describe its globalization. These notions are illustrated by an example in Section 3. Next we pass to languages which are recognized by preautomata. The Eilenberg theorem is generalized and examples are given which show that the Kleene theorem does not hold for preautomata (Section 6). In the last section we compare the preautomata with other classes of machines.

We use the notation $\varphi : A \dashrightarrow B$ for a partial mapping from A to B (unlike to a full mapping $A \to B$). If $\varphi(a)$ is not defined for $a \in A$ we write $\varphi(a) = \emptyset$. The free semigroup on the alphabet Σ is denoted by Σ^+, the free monoid on Σ by Σ^* and its identity element by ε. We use right preactions, as it is customary in Automata Theory to deal with right actions.

1. Preactions

We shall need some known information about partial actions (preactions) of monoids which we recall in this section.

Definition 1. [7] *Let M be a monoid with the identity e, X a set.* **A preaction** *M on X is a partial mapping $X \times M \dashrightarrow X : (x, a) \mapsto xa$, such that*

$$\forall x \in X \quad xe = x,$$

$$\forall a, b \in M \;\; \forall x \in X \;\; xa \neq \emptyset \;\&\; (xa)b \neq \emptyset \implies x(ab) \neq \emptyset \;\&\; (xa)b = x(ab),$$

$$\forall a, b \in M \;\; \forall x \in X \;\; xa \neq \emptyset \;\&\; x(ab) \neq \emptyset \implies (xa)b \neq \emptyset \;\&\; (xa)b = x(ab).$$

The preactions of the given monoid M form a category $\mathcal{PA}M$: its objects are sets with preactions given on them, and morphisms are such mappings $\varphi : X \to Y$ that

$$\forall a \in M \; \forall x \in X \;\; xa \neq \emptyset \implies \varphi(x)a \neq \emptyset \;\&\; \varphi(xa) = \varphi(x)a.$$

Preactions appear in the following situation. Suppose that a full action of a monoid M is given on a set Y and $X \subset Y$. Then the restriction of the action to X is a preaction.

Conversely, let $X \times M \dashrightarrow X$ be a preaction and $Y \supset X$.

Definition 2. *A full action $Y \times M \to Y$ is called a* **globalization** *if its restriction on X coincides with the given preaction.*

The following construction gives a globalization for any preaction $X \times M \dashrightarrow X$. Define on the set $X \times M$ a relation $\vdash$:

$$(x, ab) \vdash (xa, b) \iff xa \neq \emptyset.$$

Let $\simeq$ is the equivalence generated by $\vdash$ and $Y = (X \times M)/\simeq$. Denote by $[x,a]$ the $\simeq$-class containing the pair (x,a). Set $[x,a]b = [x,ab]$ for $[x,a] \in Y,\ b \in M$. This defines a full action on Y.

Theorem 1. [7] *The above defined action $Y \times M \to Y$ is a globalization of the preaction $X \times M \dashrightarrow X$; the mapping $\alpha : X \to Y : x \mapsto [x,e]$ is an injective morphism, and any morphism of the preaction M on X in a full action of M can be factored through α.*

2. Preautomata

We will use the definition of an automaton in the following form (in this section the condition of finiteness is ignored):

Definition 3. *Let X be a set and Σ^* the free monoid over an alphabet Σ. An* **automaton** *is a triple (Σ, X, δ^*) where $\delta^* : X \times \Sigma^* \to X$ is a mapping such that*

$$\forall x \in X \quad \delta^*(x,\varepsilon) = x, \tag{1}$$

$$\forall u,v \in \Sigma^* \ \forall x \in X \quad \delta^*(x,uv) = \delta^*(\delta^*(x,u,),v). \tag{2}$$

The main object of this article is a more general concept:

Definition 4. *A* **preautomaton** *is such a triple (Σ, X, δ^*), where $\delta^* : X \times \Sigma^* \dashrightarrow X$ is a partial mapping, that*

a) the above condition (1) is satisfied;

b) if $\delta^*(x,u) \neq \emptyset$ and $\delta^*(\delta^*(x,u),v) \neq \emptyset$, then $\delta^*(x,uv) \neq \emptyset$ and

$$\delta^*(x,uv) = \delta^*(\delta^*(x,u),v); \tag{3}$$

c) if $\delta^*(x,u) \neq \emptyset$ and $\delta^*(x,uv) \neq \emptyset$, then $\delta^*(\delta^*(x,u),v) \neq \emptyset$ and also the equation (3) is fulfilled.

Clearly, preautomata correspond to preactions of the free monoid Σ^*. It will be convenient to omit the symbol δ^*, i.e. to write xu instead of $\delta^*(x,u)$, and denote the preautomaton by (Σ, X).

Theorem 1 enables us to associate to the preautomaton $\mathcal{M} = (\Sigma, X)$ an automaton $\mathcal{M}_{\mathrm{gl}} = (\Sigma, Y)$, where $Y = (X \times \Sigma^*)/\simeq$ is the set constructed in Section 1. We call $\mathcal{M}_{\mathrm{gl}}$ a **globalization** of $\mathcal{M}$.

Using the fact that the monoid Σ^* is free, the description of $\simeq$ can be simplified:

Theorem 2. *Define the following relation $\approx$ on the set $X \times \Sigma^*$:*

$$(x,a) \approx (y,b) \Longleftrightarrow \exists\ a', b', p \in \Sigma^* : a = a'p,\ \ b = b'p,\ \ xa' = yb' \neq \emptyset.$$

Then $\approx$ coincides with $\simeq$.

Proof. If $(x,a) \approx (y,b)$ then $(x,a'p) \vdash (xa',p) = (yb',p)$ and $(y,b'p) \vdash (yb',p)$, whence $\approx \subseteq \simeq$. Let us prove the converse inclusion.

Obviously, $\approx$ is reflexive and symmetric. Let us check its transitivity.

Let $(x,a) \approx (y,b) \approx (z,d)$, $a = a'p$, $b = b'p = c'q$, $d = d'q$ and $xa' = yb' \neq \emptyset$, $yc' = zd' \neq \emptyset$. Since Σ^* is free, either p divides q or vice versa. Let, say, $p = uq$. Then $c' = b'u$.

Since $yb' \neq \emptyset$, $y(b'u) \neq \emptyset$ then $(xa')u = (yb')u = y(b'u) \neq \emptyset$. Hence $(xa')u = x(a'u)$, as $xa' \neq \emptyset$. On the other hand $a'p = (a'u)q$, and we obtain $(x,a'p) = (x,(a'u)q) \approx (z,d'q)$, since $(xa')u = (yb')u = y(b'u) = yc' = zd'$. Thus, $\approx$ is an equivalence, and $\approx \supset \simeq$ as clearly $\approx \supset \vdash$. $\square$

Corollary 1. *Every $\simeq$-class can be uniquely written either in the form $[x,w]$, where $x \in X$, $w = a_1 \dots a_n \in \Sigma^*$ and $x(a_1 \dots a_i) = \emptyset$ for any $i, 1 \leq i \leq n$, or in the form $[x,\varepsilon]$ ($n = 0$).*

Theorems 1 and 2 allow to give the following interpretation of a finite preautomaton. Suppose that a system with a possibly infinite set of states is given, and only a finite subset of its states is accessible to our observation. Then the preautomaton, obtained by restriction of the system to the set of observable states, can be considered as a model of the observable part of the system.

3. Example

Consider the infinite automaton $\mathcal{N} = (\Sigma, \mathbb{Z})$ for which $\Sigma = \{a,b\}$ is a two-lettered alphabet, $\mathbb{Z}$ is the set of integers and the transition function is defined by the equalities $n \cdot a = n+1$, $n \cdot b = n-1$. Observe that this is the natural action of Σ^* on its factor-monoid by the congruence generated by the relation $ab = \varepsilon$, and this factor-monoid is isomorphic to the infinite cyclic group. Set $X = \{0,1\} \subset \mathbb{Z}$. Then the restriction of the action of Σ^* on X gives a preautomaton $\mathcal{M} = (\Sigma, X)$. In order to describe the corresponding preaction, we denote the length of a word w by $|w|$, the number of entries of the letter a in w by $|w|_a$, and set $\|w\| = |w|_a - |w|_b$. Then

$$0 \cdot w = \begin{cases} 0 & \text{if } \|w\| = 0, \\ 1 & \text{if } \|w\| = 1, \\ \emptyset & \text{otherwise,} \end{cases} \qquad 1 \cdot w = \begin{cases} 0 & \text{if } \|w\| = -1, \\ 1 & \text{if } \|w\| = 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

Corollary 1 allows to describe the globalization $\mathcal{M}_{\mathrm{gl}} = (\Sigma, Y)$. A word $w = a^{\alpha_1}b^{\beta_1}\ldots a^{\alpha_n}b^{\beta_n}$ will be called *1-simple* if $\|a^{\alpha_1}b^{\beta_1}\ldots a^{\alpha_i}b^{\beta_i}\| > 0$ for all $i \geq 1$ (this implies, in particular, that $\alpha_1 > 0$). The word ε will be considered 1-simple too. Similarly, a word $w = b^{\beta_1}a^{\alpha_1}\ldots b^{\beta_n}a^{\alpha_n}$ (and also the word ε) is *0-simple* if $\|b^{\beta_1}a^{\alpha_1}\ldots b^{\beta_i}a^{\alpha_i}\| < 0$ for all $i \geq 1$. It is easy to see that every $\simeq$-class has form $[0, w]$ or $[1, w]$, where the word w is 0-simple or 1-simple, respectively.

Note that the preautomaton $\mathcal{M} = (\Sigma, X)$ cannot be considered as a restriction of a finite automaton. Indeed, let $\mathcal{P} = (\Sigma, W)$ be such an automaton with a transition function $\delta : W \times \Sigma \to W$ and $X \subset W$. Since W is finite, $\delta(0, b^k) = \delta(0, b^m)$ for some distinct k, m. Suppose $k < m$. As $\|b^k a^k\| = 0$ then $\delta(\delta(0, b^k), a^k) = \delta(0, b^k a^k) = 0 \cdot (b^k a^k) = 0$. Further, $\delta(0, b^m a^k) = \delta(\delta(0, b^m), a^k) = \delta(\delta(0, b^k), a^k) = 0$, and since our preautomaton is a restriction of $\mathcal{P}$, then $0 \cdot (b^m a^k) = 0$. This contradicts the inequality $\|b^m a^k\| < 0$.

In addition, it is possible to obtain from Theorem 2 the description of the semigroup of $\mathcal{M}_{\mathrm{gl}} = (\Sigma, Y)$. We recall that the semigroup of an automaton is the factor-monoid obtained from Σ^* by identification of the words equally acting of the states.

Proposition 1. *The semigroup of $\mathcal{M}_{\mathrm{gl}} = (\Sigma, Y)$ coincides with $\Sigma^* = \{a, b\}^*$.*

Proof. Suppose that the words $u, v \in \Sigma^*$, $u \neq v$, act equally on all states of $\mathcal{M}_{\mathrm{gl}} = (\Sigma, Y)$. In particular, $[0, b^\beta]u = [0, b^\beta]v$ for all $\beta > 0$. As $b^\beta u = b^\beta v$, then by Theorem 2 $u = u'w$, $v = v'w$ and $0 \cdot b^\beta u' = 0 \cdot b^\beta v' \neq \emptyset$. Hence,

$$\|u'\| = \|v'\| = \left\{ \begin{array}{l} \beta \\ \beta + 1 \end{array} \right. \geq \beta.$$

But $\|u'\| = |u'|_a - |u'|_b \leq |u'| \leq |u|$ is a bounded quantity in contrary with arbitrariness of β. □

Remark. It is easy to see that the semigroup of the original automaton $\mathcal{N} = (\Sigma, Z)$ is isomorphic to $\mathbb{Z}$.

4. Recognizability

In what follows we will consider only **finite** preautomata, i. e. preautomata $\mathcal{M} = (\Sigma, X)$ such that $|X| < \infty$. A preautomaton $\mathcal{M} = (\Sigma, X)$, in which **an initial state** $x_0 \in X$ and **a terminal subset** $T \subset X$ are chosen, will be called **a preacceptor** and denoted by $\mathcal{M} = (\Sigma, X, x_0, T)$.

As well as in the classical situation, we call a language $L \subset \Sigma^*$ **recognizable** if there is a preacceptor $\mathcal{M} = (\Sigma, X, x_0, T)$ for which

$$L = \{w \in \Sigma^* \mid \emptyset \neq x_0 w \in T\}.$$

In what follows recognizability will be understood in this sense.

The following assertion generalizes the Myhill–Nerode theorem and gives an algebraic characterization of recognizable languages:

Theorem 3. *A language $L \subset \Sigma^*$ is recognized by a preacceptor if and only if L is the union of a finite number of classes of some right congruence on Σ^*.*

Proof. Suppose that L is recognized by a (finite) preacceptor $\mathcal{M} = (\Sigma, X, x_0, T)$. Consider its globalization $\mathcal{M}_{\mathrm{gl}} = (\Sigma, Y)$. It follows from Corollary 1 that L is recognized by the acceptor

$$\mathcal{N} = (\Sigma, [x_0, \varepsilon]\Sigma^*, [x_0, \varepsilon], [T, \varepsilon])$$

where $[T, \varepsilon] = \{[t, \varepsilon] \mid t \in T\}$, $[x_0, \varepsilon]\Sigma^* \subseteq Y$. The relation

$$\rho = \{(u, v) \in \Sigma^* \times \Sigma^* \mid [x_0, \varepsilon]u = [x_0, \varepsilon]v\}$$

is a right congruence, L is a union of ρ-classes and the number of these classes does not exceed $|[T, \varepsilon]| = |T| < \infty$.

Conversely, let ρ be a right congruence on Σ^* and L be the union of a finite number (say, n) of ρ-classes $C_1, \ldots, C_n$.

In the (infinite) automaton $\mathcal{K} = (\Sigma, \Sigma^*/\rho)$ we choose the ρ-class E containing ε as the initial state, and $T = \{C_1, \ldots, C_n\} \subset \Sigma^*/\rho$ as the terminal subset. Then $(\Sigma, T \cup \{E\}, E, T)$ is a finite preacceptor which is a restriction of $\mathcal{K}$ and recognizes L. $\square$

Corollary 2. *Given languages $L, M \subset \Sigma^*$ which are recognized by preacceptors, then $L \cap M$ is also recognizable.*

Proof. Let L and M are finite unions of classes of right congruences λ and μ respectively. Then $L \cap M$ is a union of a finite number of classes of the congruence $\lambda \cap \mu$. $\square$

Corollary 3. *For the single-letter alphabet $\Sigma = \{a\}$ a language $L \subset \Sigma^*$ is recognizable by a preacceptor exactly when L is recognizable by an acceptor.*

Proof. The semigroup $\Sigma^+ = \{a\}^+$ is the unique infinite monogenic semigroup up to isomorphism [6]. As it is commutative, all its right congruences are two-sided ones. Factor-semigroups by these congruences (except the trivial one) are finite. Therefore if a language L is recognized by a preacceptor, but not by an acceptor, it should be a union of a finite number of classes of the trivial congruence, i. e. L is a finite subset in Σ^+, hence L is regular. □

Example 1. The language $L = \{a^n b^n \mid n > 0\}$ over the alphabet $\Sigma = \{a, b\}$ is a class of the right syntactic congruence [6]. Therefore it is recognized by a preacceptor. As it is well-known [6], L is not recognized by any finite acceptor.

Example 2. For an arbitrary finite Σ each ideal of the monoid Σ^* is recognizable, since it is an element of the Rees factor-semigroup.

5. Minimization of a preacceptor

The study of recognizability of languages by preacceptors leads to the notion of the syntactic equivalence of preacceptors.

Definition 5. *Preacceptors $\mathcal{M}_1$ and $\mathcal{M}_2$ over the same alphabet Σ are called* **syntactically equivalent** *if they recognize the same language.*

As in the theory of acceptors, a question arises how to find in a class of syntactically equivalent preacceptors a preacceptor whose set of states has minimal cardinality. The first step is given by the following lemma.

Lemma 1. *Let $\mathcal{M} = (\Sigma, X, x_0, T)$ be a preacceptor, $Y = \{x_0\} \cup T$, $\mathcal{M}_0 = (\Sigma, Y, x_0, T)$ the restriction of $\mathcal{M}$ on Y. Then $\mathcal{M}_0$ is syntactically equivalent to $\mathcal{M}$.*

Proof. is obvious. □

Thanks to Lemma 1, in the study of language recognition by preacceptors we can restrict the set of states including only the initial and terminal states.

We recall from [6] that the right syntactic congruence on Σ^* of a language $L \subset \Sigma^*$ is the relation $\equiv_L$ defined by

$$w_1 \equiv_L w_2 \Longleftrightarrow \forall u \in \Sigma^* \ (w_1 u \in L \Leftrightarrow w_2 u \in L)$$

The language L is the union of some classes of this relation; moreover, any right congruence, such that L is a union of its classes, is contained in $\equiv_L$. This property allows us to reformulate Theorem 3 as follows:

Lemma 2. *A language $L \subset \Sigma^*$ is recognized by a (finite) preacceptor if and only if it is the union of a finite number of classes of its right syntactic congruence.*

We will denote the $\equiv_L$-class containing $u \in \Sigma^*$ by $[u]_{\equiv_L}$. Fix a set of representatives $u_1, \ldots, u_k \in L$ of the $\equiv_L$-classes of L.

For the set $X_{\equiv_L} = \{[\varepsilon]_{\equiv_L}, [u_1]_{\equiv_L}, \ldots, [u_k]_{\equiv_L}\} \subset \Sigma^* / \equiv_L$ we define the partially defined map $\delta^*\colon X_{\equiv_L} \times \Sigma^* \dashrightarrow X_{\equiv_L}$ as follows:

$$\forall\, [u]_{\equiv_L} \in X_{\equiv_L} \quad \delta^*([u]_{\equiv_L}, \varepsilon) = [u]_{\equiv_L}, \tag{4}$$

$$\forall\, [u]_{\equiv_L} \in X_{\equiv_L}\ \forall\, w \in \Sigma^+ \quad uw \notin L \Rightarrow \delta^*([u]_{\equiv_L}, w) = \emptyset, \tag{5}$$

$$\forall\, [u]_{\equiv_L} \in X_{\equiv_L} \quad \forall\, w \in \Sigma^+ \quad uw \in L \Rightarrow \delta^*([u]_{\equiv_L}, w) = [uw]_{\equiv_L}. \tag{6}$$

Theorem 4. *Let $L \subset \Sigma^*$ is a finite union of $\equiv_L$-classes. Then the preacceptor*

$$\mathcal{M}_{\equiv_L} = (\Sigma, X_{\equiv_L}, [\varepsilon]_{\equiv_L}, \{[u]_{\equiv_L} \mid u \in L\})$$

with the partial mapping δ^*, defined by (4)–(6), recognizes L.

Moreover, $\mathcal{M}_{\equiv_L}$ has the minimal set of states among all finite preacceptors recognizing L.

Proof. First of all we check that δ^* defines a preaction. Indeed, condition a) of Definition 4 is fulfilled by virtue of (4), and conditions b) and c) follow from (5) and (6). Thus, $\mathcal{M}_{\equiv_L}$ is a preacceptor which recognizes L.

Now, let $\mathcal{M} = (\Sigma, X, x_0, T)$ be a preacceptor recognizing L. Define a mapping $f\colon X_{\equiv_L} \to X$ by the formulas

$$f([\varepsilon]_{\equiv_L}) = x_0,$$

$$f([u_i]_{\equiv_L}) = x_0 \cdot u_i, \quad i = 1, \ldots, k.$$

As $u_i \in L$, $x_0 \cdot u_i \in T$.

We wish to show that f is injective. Assume that $x_0 \cdot u_i = x_0 \cdot u_j$ for some $1 \le i \ne j \le k$. Suppose that $u_i w \in L$, so $\emptyset \ne x_0 \cdot (u_i w) \in T$. Since $u_i \in L$, then $\emptyset \ne (x_0 \cdot u_i) \cdot w = x_0 \cdot (u_i w) \in T$ by the definition of a preautomaton. The assumption $x_0 \cdot u_i = x_0 \cdot u_j$ implies $\emptyset \ne x_0 \cdot (u_j w) = (x_0 \cdot u_j) \cdot w \in T$, i. e. $u_j w \in L$. Thus $u_i w \in L$ yields $u_j w \in L$, and similarly $u_j w \in L \Longrightarrow u_i w \in L$. Hence $u_i \equiv_L u_j$, contradicting $i \ne j$. Thus f is injective and $|X_{\equiv_L}| \le |X|$. □

6. Prefix decomposition

In this section Theorem VI.4.1 from [1] is generalized to finite preautomata.

Lemma 3. *Suppose that L is a language over an alphabet Σ which is recognized by a preacceptor $\mathcal{M} = (\Sigma, X, x_0, T)$ and let L_y $(y \in T)$ be languages such that L_y is recognized by a preacceptor $\mathcal{M}_y = (\Sigma, X, x_0, T_y = \{y\})$. Then $L = \coprod_{y \in T} L_y$ (disjoint union).*

Proof. $w \in L$ if and only if $x_0 w = y$ for some $y \in T$, i.e. $w \in L_y$. □

We remind that a (possibly, empty) language $L \subset \Sigma^*$ is called **a prefix code** [6] if $u, uv \in L$ implies $v = \varepsilon$. The language $\{\varepsilon\}$ is also considered as a prefix code (note that if a prefix code contains ε then it is equal $\{\varepsilon\}$). It is well-known and easily verified that the monoid $S = L^*$, generated by a prefix code L, is free. Such a monoid is called **unitary** and is characterized by the property:

$$u, uv \in S \Longrightarrow v \in S. \tag{7}$$

The following lemma generalizes (7):

Lemma 4. *Let H and C be prefix codes over an alphabet Σ. Then*

$$u, uv \in HC^* \Longrightarrow v \in C^*.$$

Proof. Suppose that

$$u = hc_1 \dots c_m,\ uv = h'c'_1 \dots c'_n\ (h, h' \in H, c_i, c'_j \in C).$$

Then $hc_1 \dots c_m v = h'c'_1 \dots c'_n$. Since H is a prefix code, then $h = h'$, and hence $c_1 \dots c_m v = c'_1 \dots c'_n$. This implies $c_i = c'_i$ $(1 \le i \le m \le n)$ and $v = c'_{m+1} \dots c'_n \in C^*$. □

It is interesting to note that the property of prefix codes to generate free monoids can be generalized as follows:

Proposition 2. [1] *Let $P_1, \dots, P_n$ $(n \ge 1)$ be prefix codes over Σ and $w \in P_1 \cdot \ldots \cdot P_n$. Then the decomposition $w = w_1 \dots w_n$, where $w_i \in P_i$ for $i = 1, \dots, n$, is unique.*

Proof. We use induction on n. For $n = 1$ the statement is evident. Suppose that it is true for $k < n$. Let $w = w'_1 \dots w'_n = w''_1 \dots w''_n$ be two representations of a word $w \in P_1 \cdot \ldots \cdot P_n$ where $w'_i, w''_i \in P_i$ for $i = 1, \dots, n$. Denote $u' = w'_2 \dots w'_n$, $u'' = w''_2 \dots w''_n$ and assume that $|u'| \ne |u''|$, for example, $|u'| < |u''|$. Then $|w'_1| > |w''_1|$, so w''_1 is a prefix of w'_1, which is impossible for a prefix code.

Hence, $|u'| = |u''|$. Then $u' = u''$ as suffixes of w of the same length, and $w'_1 = w''_1$ as prefixes. By induction the equality $u' = u''$ implies $w'_2 = w''_2, \dots, w'_n = w''_n$. □

Definition 6. *A language $L \in \Sigma^*$, such that $L = HC^*$ where H and C are prefix codes over Σ, will be called a p-***language**.

[1] Proposition 2 will not be used in this article.

Remark. Notice that if $H = \emptyset$ then $HC^* = \emptyset$.

Theorem 5. *A language L over an alphabet Σ is recognized by some preacceptor $\mathcal{M} = (\Sigma, X, x_0, \{y\})$ with a single terminal state if and only if L is a p-language.*

Proof. Let L be a p-language. We consider the possible cases.

1) $L_1 = \emptyset$. Then L_1 is recognized by the preacceptor

$$\mathcal{M} = (\Sigma, \{x_0, y\}, \{x_0\}, \{y\})$$

where the action is trivial:

$$x_0 w = x_0, \; yw = y.$$

2) $L_2 = C^*$. Then L_2 is recognized by the preacceptor

$$\mathcal{M} = (\Sigma, \{x_0\}, \{x_0\}, \{x_0\})$$

where preaction It is given by formulas:

$$x_0 w = \begin{cases} x_0 & \text{if } w \in C^*, \\ \emptyset & \text{if } w \notin C^*. \end{cases}$$

The fact that the above is a preaction, is verified using (7).

3) $L_3 = HC^*$ with $H \neq \{\varepsilon\}$. Then L_3 is recognized by the preacceptor $\mathcal{M} = (\Sigma, \{x_0, y\}, \{x_0\}, \{y\})$ with the preaction:

$$x_0 w = \begin{cases} y, & \text{if } w \in HC^*, \\ x_0, & \text{if } w = \varepsilon, \\ \emptyset, & \text{if } w \notin HC^* \cup \{\varepsilon\}, \end{cases} \qquad yw = \begin{cases} y & \text{if } w \in C^*, \\ \emptyset & \text{if } w \notin C^*. \end{cases}$$

We need to show that the above formulas define a preautomaton. For the action of an input word on the state y the fulfilment of the conditions of Definition 4 follows from (7). As to the action on x_0, the condition b) of Definition 4 is immediate, and in order to see c) suppose that $x_0 u \neq \emptyset$ and $x_0(uv) \neq \emptyset$. If $u = \varepsilon$, the condition c) is obvious. Otherwise $u, uv \in HC^*$ and by Lemma 4, $v \in C^*$.

Now we prove that if a language L is recognized by a preacceptor of the form $\mathcal{M} = (\Sigma, X, x_0, \{y\})$ then it is a p-language. For every word $w \in \Sigma^*$ denote by $\mathrm{Px}(w)$ the set of all proper prefixes of w. For $K \subset \Sigma^*$ write $\mathrm{Px}(K) = \{\mathrm{Px}(w) \mid w \in K\}$. Set

$$H = \{w \in \Sigma^* \mid x_0 w = y \;\&\; \forall u \in \mathrm{Px}(w) \; x_0 u \neq y\}.$$

and similarly,

$$C = \{w \in \Sigma^* \mid yw = y \ \&\ \forall u \in \mathrm{Px}(w)\ yu \neq y\}.$$

By construction H and C are prefix codes and, moreover, $L \supset HC^*$ since $x_0 w = y$ for any word $w \in HC^*$. It remains to show that $L \subset HC^*$.

Let $w \in L$, i.e. $x_0 w = y$. First assume that $x_0 = y$. Write $w = uv$, where u is the least prefix of w for which $x_0 u = x_0$ (and therefore, $u \in H = C$). By definition of a preaction $\emptyset \neq (x_0 u)v = x_0 w$, whence $x_0 v = x_0$, and consequently $v \in C^*$. Thus we obtain $w = uv \in HC^* = C^*$.

If $x_0 \neq y$, then write $w = uv$, where u is the prefix of the least length, such that $x_0 u = y$. It follows that $u \in H$ and $yv = y$. As above, $v \in C^*$, whence $w \in HC^*$. □

Theorem 5 allows us to transfer to preautomata the concept of a prefix decomposition ([1], Theorem VI.4.1):

Corollary 4. *If a language L over the alphabet Σ is recognized by some preacceptor then L decomposes into a disjoint union*

$$L = \coprod_i H_i C_i^*,$$

where H_i, C_i are prefix codes over Σ.

Proof. The corollary follows from Theorem 5 and Lemma 3. □

The converse statement does not hold:

Example 3. It is directly seen that the language $H_1 = \{a\}^+$ is recognized (even by a finite acceptor), and the language $H_2 = \{a^n b^n \mid n \geq 1\}$ is a prefix code. Take $C_1 = C_2 = \{\varepsilon\}$. The union $H = H_1 \cup H_2 = H_1 C_1^* \cup H_2 C_2^*$ is not recognized by any preacceptor. Indeed, let $\equiv_H$ be the right syntactic congruence of H, and take $x \in \Sigma^*$. Since

$$a^n x \in H \iff x = b^n \text{ or } x \in H_1 \cup \{\varepsilon\},$$

the words a^n form one-element $\equiv_H$-classes, and by Theorem 3 the language H is non-recognizable.

The given example shows also that the set of languages, recognized by finite preacceptors, is not union-closed. Moreover, it is not closed with respect to the other Kleene's operations, i.e. the product and the iteration:

Example 4. Let H_1 and H_2 be the same as in Example 3. By Theorem 5 the language H_2^* is recognizable. Consider the product $K = H_2^* H_1$. Then $a^n \in K$ for $n > 0$ and

$$\{x \in \Sigma^* \mid a^n x \in K\} \cap \{b\}^+ = \{b^n\}.$$

Therefore all words of the form a^n are pairwise non-equivalent with respect to the right congruence $\equiv_K$. As in Example 3, it follows that K is non-recognized.

Example 5. The language $L = H_2 \cup \{a\}$ is recognized by a preacceptor, since it consists of two $\equiv_L$-classes H_2 and $\{a\}$. Consider the iteration $M = L^*$. It is easy to see that $a^m b^n \in M$ iff $m \geq n$. Hence,

$$\{x \in \Sigma^* \mid a^m x \in M\} \cap \{b\}^+ = \{b^k \mid k \leq m\}.$$

As above, the words a^m are pairwise non-equivalent with respect to $\equiv_M$. Therefore, $M = L^*$ is not recognized by any preacceptor.

Conclusion

The obtained results allow to clarify relations between the class of finite preautomata FPA and other types of machines.

Obviously, FPA includes the class of finite automata FA. This inclusion is strict since the language from Example 1 is not recognized by any finite automaton. On the other hand, it is known [6, § 9.1] that the prefix language $\{a^n b^n c^n \mid n \geq 1\}$ is not context-free. According to Theorem 5 it means that FPA is not contained in the class of automata with stack memory FSA.

At the same time, FPA does not include the class of Turing machines TM. Indeed, the language $\{a^{n^2} \mid n \geq 1\}$ is not recognized by a finite automaton [5, ex. 4.1.2], and by Corollary 3 it is not recognized also by a preautomaton.

On the other hand, the cardinal of the class of all prefix codes over some finite alphabet equals to the cardinal of continuum, whereas the class of a recursively enumerable languages over a finite alphabet is countable. It follows that TM does not contain FPA.

References

[1] S. Eilenberg, *Automata, Languages, and Machines*, vol. B., Academic Press, 1976.

[2] R. Exel, *Partial actions of groups and actions of semigroups*, Proc. Amer. Math. Soc., **126**(1998), pp.3481-3494.

[3] W. M. L. Holcombe, *Algebraic Automata Theory*, Cambridge Univ. Press, 1982.

[4] C. Hollings, *Partial actions of monoids*, Semigroup Forum, **75**(2007), pp.293-316.

[5] J. E. Hopcroft, R. Motwani, J. D. Ullman, *Introduction to Automata Theory, Languages and Computation*, 2nd edition, Addison Wesley, 2001.

[6] G. Lallement, *Semigroups and Combinatorial Applications*, Pure and Applied Mathematics Series, Wiley, New York, 1979.

[7] M. Megrelishvili, L. Schröder, *Globalization of confluent partial actions on topological and metric spaces*, Topology and its Appl., **145**(2004), pp.119-145.

Contact information

M. Dokuchaev Instituto de Matemática e Estatística Universidade de São Paulo, Rua do Matão, 1010, CEP 05508-090, São Paulo, SP, Brazil
E-Mail: dokucha@ime.usp.br

B. Novikov Kharkov National University, Svobody sq., 4, 61077, Kharkov, Ukraine
E-Mail: novikov@univer.kharkov.ua

G. Zholtkevych Kharkov National University, Svobody sq., 4, 61077, Kharkov, Ukraine
E-Mail: g.zholtkevych@gmail.com

Received by the editors: 13.04.2011
and in final form 05.05.2011.