On Varieties of Literally Idempotent Languages

Ondřej Klíma and Libor Polák *

Department of Mathematics, Masaryk University
Janáčkovo nám 2a, 662 95 Brno, Czech Republic
http://www.math.muni.cz/~{klíma, polak}

Abstract. A language $L \subseteq A^*$ is literally idempotent in case that $ua^2v \in L$ if and only if $uav \in L$ for each $u, v \in A^*$, $a \in A$. Such classes result naturally by taking all literally idempotent languages in a classical (positive) variety or by considering a certain closure operator. We initiate their systematic study. Various classes of such languages can be characterized using syntactic methods. A starting example is the class of all finite unions of $B_1^* B_2^* \dots B_k^*$ where $B_1, \dots, B_k$ are subsets of a given alphabet A.

MSC 2000 Classification: 68Q45 Formal languages and automata

1 Introduction

Papers by Straubing [10] on $\mathbb{C}$-varieties and Ésik and el. [4], [5] on literal varieties of languages enable us to consider new significant classes of languages. Due to the result by Kunc [7] we also have equational logic for those classes.

(Positive) varieties of languages corresponding to pseudovarieties of (ordered) idempotent semigroups/monoids are not very important from the point of language theory. This is far from being the case for languages corresponding to pseudovarieties of literally idempotent homomorphisms.

Most of our classes result by considering intersections of well-known classical (positive) varieties with literally idempotent languages. Our new classes nicely fit into the table in Section 8 by Pin [9]. We characterize languages from certain classes of languages in various ways. More precisely we describe the languages which are literally idempotent and which belong to the level 1/2, level 1, level 3/2 respectively. We also consider other interesting classes of languages, e.g. languages which are finite unions of the languages of the form $B_0^* B_1^* \dots B_k^*$, where $k \in \mathbb{N}_0$ and $B_0, \dots, B_k$ are subsets of a given alphabet.

We also describe all literally idempotent languages over two element alphabet and we present canonical forms of the corresponding regular expressions.

Notice that the motivation for studying literally idempotent languages also comes from the linear temporal logic. The formulas of LTL without the "next" operator determine literally idempotent languages. We give a logical characterization of languages from two of our classes.

* Both authors acknowledge the support of the Grant no. 201/06/0936 of the Grant Agency of the Czech Republic.

The paper is organized as follows. In Section 2 we recall known results and techniques obtained by syntactic methods. In Section 3 we present several new classes of language. In Section 4 introduce literally idempotent languages and their basic properties. In Section 5 we prove results concerning intersections of literally idempotent languages with well-known classes (level 1/2, 1, 3/2, right-trivial languages, finite languages). In Section 6 we comment on literally idempotent languages over a two element alphabet. The last section contains several remarks dealing with the relationship to the linear temporal logic.

2 Preliminaries

Valuable treatments on syntactic methods in language theory are books by Almeida [1], Pin [8] and his chapter [9].

Let $\mathcal{M}$ (resp. $\mathcal{O}$) be the class of all surjective homomorphisms from free monoids over non-empty finite sets onto finite (ordered) monoids. A class $\mathcal{V} \subseteq \mathcal{M}$ is a *literal pseudovariety* if it is closed with respect to the homomorphic images, literal substructures and products of finite families – see Ésik and el. [4], [5] or Straubing [10] for a more general notion of a $\mathbb{C}$-pseudovariety. Similarly, we define the literal pseudovarieties in the ordered case.

Let $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Let I_n, for $n \in \mathbb{N}$, be the set of all n-ary implicit operations for the class of finite monoids – see [1]. We write $\pi^M : M^n \to M$ for the realization of $\pi \in I_n$ on a finite monoid M. A pseudoidentity $\pi = \rho$, where $\pi, \rho \in I_n$, is *literally* satisfied in

$$(\phi : A^* \to M) \in \mathcal{M}$$

$$\text{if } \ (\ \forall\, a_1, \dots, a_n \in A\ )\ \pi^M(\phi(a_1), \dots, \phi(a_n)) = \rho^M(\phi(a_1), \dots, \phi(a_n))\ .$$

We write $\phi \models_{\mathfrak{L}} \pi = \rho$ in this case.

Similarly, a pseudoinequality $\pi \leq \rho$, where $\pi, \rho \in I_n$, is *literally* satisfied in

$$(\phi : A^* \to (M, \leq)) \in \mathcal{O}$$

$$\text{if } (\ \forall\, a_1, \dots, a_n \in A\ )\ \pi^M(\phi(a_1), \dots, \phi(a_n)) \leq \rho^M(\phi(a_1), \dots, \phi(a_n))\ .$$

We write $\phi \models_{\mathfrak{L}} \pi \leq \rho$ in this case.

Usually we fix an alphabet $\Sigma = \{x_1, \dots, x_n\}$ of variables and we identify a word $u = x_{i_1} \dots x_{i_k} \in \Sigma^*$ with the implicit operation $u^M(a_1, \dots, a_n) = a_{i_1} \dots a_{i_k}$, where $M \in \mathcal{M},\ a_1, \dots, a_n \in M$. Examples of implicit operations which are not of this form are u^ω, for $u \in \Sigma^+$. We define

$$((x_{i_1} \dots x_{i_k})^\omega)^M(a_1, \dots, a_n) = a^\omega\ ,$$

where $a = a_{i_1} \dots a_{i_k}$ and a^ω is the unique idempotent in the set $\{a, a^2, a^3, \dots\}$.

A special case of the main result of Kunc [7] follows.

Result 1 (Kunc) *The literal pseudovarieties of homomorphisms onto finite monoids are exactly the subclasses of $\mathcal{M}$ defined by the literal satisfaction of sets of pseudoidentities.*

One can expect an analogous result in the ordered case – we do not need it here, we only support it by examples.

By a *quotient* of $L \subseteq A^*$ we mean any set $u^{-1}Lv^{-1} = \{ w \in A^* \mid uwv \in L \}$ where $u, v \in A^*$.

A *class* of (regular) languages is an operator $\mathscr{V}$ assigning to each non-empty finite set A a set $\mathscr{V}(A)$ of regular languages over the alphabet A.

Such a class is a *positive variety* if
(0) for each A, we have $\emptyset$, $A^* \in \mathscr{V}(A)$,
(i) each $\mathscr{V}(A)$ is closed with respect to finite unions, finite intersections and quotients, and
(ii) for each non-empty finite sets A and B and a homomorphism $f : B^* \rightarrow A^*$, $K \in \mathscr{V}(A)$ implies $f^{-1}(K) \in \mathscr{V}(B)$.

Adding the condition
(iii) each $\mathscr{V}(A)$ is closed with respect to complements,
we get a *boolean* variety.

A modification of (ii) to
(ii') for each non-empty finite sets A and B and a homomorphism $f : B^* \rightarrow A^*$ with $f(B) \subseteq A$, $K \in \mathscr{V}(A)$ implies $f^{-1}(K) \in \mathscr{V}(B)$
leads to the notions of *literal* positive/boolean variety of languages.

Let $L \subseteq A^*$ be a regular language. Recall that the *syntactic congruence* $\sim_L$ on A^* is defined by

$$u \sim_L v \text{ if and only if } (\ \forall\, p, q \in A^*\ )\ (\ puq \in L \iff pvq \in L\ )\ .$$

Further, the structure $\mathsf{O}\,(L) = A^*/\sim_L$ is called the *syntactic monoid* of L and the mapping $\phi_L : A^* \rightarrow \mathsf{O}\,(L),\ u \mapsto u\sim_L$ is the *syntactic homomorphism*. Moreover, $\mathsf{O}\,(L)$ is implicitly ordered by

$$u\sim_L \leq\ v\sim_L \ \text{ if and only if } (\ \forall\, p, q \in A^*\ )\ (\ pvq \in L \implies puq \in L\ )\ .$$

We speak about the *ordered syntactic monoid* and the *ordered syntactic homomorphism*.

For a class $\mathscr{V}$ of languages, let

$$\mathsf{M}\,(\mathscr{V}) = \ \langle\ \{\ \phi_L : A^* \rightarrow \mathsf{O}\,(L) \mid A \text{ non-empty finite },L \in \mathscr{V}(A)\ \}\ \rangle$$

be the literal pseudovariety generated by the syntactic homomorphisms of members of $\mathscr{V}$, and conversely, for $\mathcal{V} \subseteq \mathcal{M}$,

$$\mathcal{V} \mapsto \mathsf{L}\,(\mathcal{V}),\ \text{ where } (\mathsf{L}\,(\mathcal{V}))(A) = \{\ L \subseteq A^* \mid \phi_L \in \mathcal{V}\ \} \text{ for each } A\ .$$

Result 2 (Ésik and Larsen [5], Straubing [10]) *The operators* M *and* L *are mutually inverse bijections between the classes of literal boolean varieties of languages and literal pseudovarieties of homomorphisms onto finite monoids.*

Similarly as in Result 1 one can expect an ordered version of Result 2 – we do not need it here, we only support it by examples.

We recall certain classical positive varieties of languages – see [8], [9].

(i) The languages of the level 1/2 over A are exactly finite unions of languages of the form

$$A^*a_1A^*a_2\dots a_kA^*,\ k\in\mathbb{N}_0,\ a_1,\dots,a_k\in A\ . \tag{1/2}$$

We denote this positive variety of languages by $\mathscr{V}_{1/2}$ and it is known that $L\in\mathscr{V}_{1/2}(A)$ iff ordered syntactic monoid of the language L satisfies the pseudoinequality $x\le 1$.

(ii) The languages of the level 1 over A are exactly the boolean combinations of languages of the form (1/2) We denote this variety of languages by $\mathscr{V}_1$ and it is known that $L\in\mathscr{V}_1(A)$ iff the syntactic monoid of the language L is $\mathcal{J}$-trivial, i.e. it satisfies the pseudoidentities $x^\omega=x^{\omega+1}$ and $(xy)^\omega=(yx)^\omega$.

(iii) The languages of the level 3/2 over A are exactly finite unions of

$$B_0^*a_1B_1^*a_2\dots a_kB_k^*,\ k\in\mathbb{N}_0,\ a_1,\dots,a_k\in A,\ B_0,\dots,B_k\subseteq A\ . \tag{3/2}$$

We denote this positive variety of languages by $\mathscr{V}_{3/2}$ and it is known that $L\in\mathscr{V}_{3/2}(A)$ iff the ordered syntactic monoid of the language L satisfies the pseudoinequalities $x^\omega yx^\omega\le x^\omega$ for every $x,y\in\Sigma^*$ such that $c(x)=c(y)$ ($c(x)$ is the set of all variables occurring in x).

(iv) We denote by $\mathscr{R}$ the positive variety of languages which can be written as (disjoint) finite unions of languages of the form

$$B_0^*a_1B_1^*a_2\dots a_kB_k^*,\qquad \text{where}$$

$$k\in\mathbb{N}_0,\ a_1,\dots,a_k\in A,\ B_0,\dots,B_k\subseteq A,\ a_i\notin B_{i-1}\text{ for } i=1,\dots,k\ . \tag{R}$$

The language L belongs to $\mathscr{R}$ iff its syntactic monoid is $\mathcal{R}$-trivial, i.e. it satisfies the pseudoidentity $(xy)^\omega x=(xy)^\omega$.

Finally, we consider two classes of +-languages together with the corresponding pseudovarieties of semigroups.

(v) Finite languages generate the positive variety of languages consisting of finite languages and full languages. This variety corresponds to the pseudovariety of ordered nilpotent semigroups with 0 being the greatest element. Such semigroups are characterized by the following pseudoinequalities

$$x^\omega y=x^\omega=yx^\omega,\ y\le x^\omega\ .$$

(vi) The boolean variety of languages generated by the class of all finite languages is the class consisting of finite and cofinite languages. This class corresponds to nilpotent semigroups.

3 New Natural Classes of Languages

In this paper we deal mainly with the following classes of languages (we will see in the next sections that they are literally idempotent). Observe the similarities with the classes of languages from Section 2.

(i) Finite unions of languages

$$A^*a_1A^*a_2\dots a_kA^*,\ k\in\mathbb{N}_0,\ a_1,\dots,a_k\in A,\ a_1\neq a_2\neq\dots\neq a_k\ . \qquad (L\ 1/2)$$

(ii) Finite unions of languages

$$B_1^*B_2^*\dots B_k^*,\ k\in\mathbb{N}_0,\ B_1,\dots,B_k\subseteq A\ . \qquad (L\ 1/2\ c)$$

(iii) Boolean combinations of languages of the form $(L\ 1/2)$.
(iv) Finite unions of languages of the form

$$B_0^*a_1B_1^*a_2\dots a_kB_k^*,\ k\in\mathbb{N}_0,\ a_1,\dots,a_k\in A,\ B_0,\dots,B_k\subseteq A,$$
$$a_1\neq a_2\neq\dots\neq a_k,\ a_1\in B_1,\dots,a_k\in B_k\ . \qquad (L\ 3/2)$$

(v) Finite unions of languages of the form

$$B_0^*a_1B_1^*a_2\dots a_kB_k^*,\ k\in\mathbb{N}_0,\ a_1,\dots,a_k\in A,\ B_0,\dots,B_k\subseteq A,$$
$$a_1\neq a_2\neq\dots\neq a_k,\ B_0\not\ni a_1\in B_1\not\ni a_2\in\dots\not\ni a_k\in B_k\ . \qquad (L\ R)$$

(vi) Finite unions of languages of the form

$$B_0^*a_1B_1^*a_2\dots a_kB_k^*,\ k\in\mathbb{N}_0,\ a_1,\dots,a_k\in A,\ B_0,\dots,B_k\subseteq A$$
$$a_1\neq a_2\neq\dots\neq a_k,\ a_1\in B_0\cap B_1,\dots,a_k\in B_{k-1}\cap B_k\ . \qquad (L\ E)$$

(vii) Finite languages generate the literal positive variety of languages, denoted by $\mathscr{N}_f$, consisting of finite languages and full languages. This variety corresponds to the variety of ordered monoids which result from nilpotent semigroups by adding units and which satisfy the pseudoinequality $x\leq 0$. This means $L\in\mathscr{N}_f(A)$ iff

$$\phi_L\models_{\mathfrak{L}} u^\omega x=u^\omega,\ xu^\omega=u^\omega,\ x\leq u^\omega,\ \text{for any } u\in\Sigma^+,\ x\in\Sigma\ .$$

(viii) The literal boolean variety of languages generated by the class of finite languages is the class $\mathscr{N}$ consisting of finite and cofinite languages. This class corresponds to nilpotent semigroups with the extra unit elements adjoined.

4 Literally Idempotent Languages

A regular language L over a finite non-empty alphabet A is *literally idempotent* if its syntactic homomorphism $\phi_L : A^*\to \mathsf{O}(L)$ satisfies the pseudoidentity $x^2=x$ literally, which means

$$(\ \forall\ a\in A\ )\ a^2\sim_L a\ .$$

We denote this class of languages by $\mathscr{L}$.

We can introduce a string rewriting system which is given by rewriting rules $pa^2q\to paq$ for any $a\in A$, $p,q\in A^*$. We say that a word $u\in A^*$ is a *normal form* of a word w if it satisfies the properties

$$w\to^* u \quad\text{and}\quad (\ u\to^* v \text{ implies } u=v\ )\ .$$

It is easy to see that this system is confluent and terminating. Consequently, for any word $w \in A^*$, there is an unique normal form $\overrightarrow{w} \in A^*$ of the word w. We will denote by $\sim$ the equivalence relation on A^* generating by the relation $\rightarrow^*$. In fact, this equivalence relation is a congruence of the monoid A^*.

For $L \subseteq A^*$ and $u \in A^*$, we write

$$u^{-1}L = \{\, w \in A^* \mid uw \in L \,\},\ \mathsf{D} = \{\ u^{-1}L \mid u \in A^*\ \}\ .$$

Classically, one assigns to L its (canonical) *minimal automaton*
$\mathcal{D} = (\mathsf{D}, A, \cdot, L, F)$ where D is the (finite) set of states,
$a \in A$ acts on $u^{-1}L$ by $(u^{-1}L) \cdot a = a^{-1}(u^{-1}L)$,
L is the initial state and $Q \in \mathsf{D}$ is a terminal state (i.e., an element of F) if and only if $1 \in Q$.

A complete deterministic automaton $\mathcal{A} = (Q, A, \cdot, i, T)$ is called *literally idempotent* if for each $q \in Q$ and $a \in A$ we have $q \cdot a^2 = q \cdot a$. Notice that the class of all such $(Q, A, \cdot)$'s forms a q-variety in the sense of Ésik and Ito [4].

In what follows we are interested in literal positive/boolean varieties consisting of literally idempotent languages. These varieties can be induced by classical varieties in two natural ways. At first, for a class of languages $\mathscr{C}$ we can consider the set of languages from $\mathscr{C}$ which are also literally idempotent languages, i.e. the intersection $\mathscr{C} \cap \mathscr{L}$ of the classes $\mathscr{C}$ and $\mathscr{L}$. The second possibility is to consider the following (closure) operator on languages. For any language $L \subseteq A^*$, we define

$$\overline{L} = \{\, w \in A^* \mid (\exists u \in L)(u \sim w) \,\} \text{ which is } \{\, w \in A^* \mid (\exists u \in L)(\overrightarrow{u} = \overrightarrow{w}) \,\}\ .$$

Lemma 1. *For $K, L \subseteq A^*$, we have :*
(i) $\overline{L}$ is regular whenever L is regular,
(ii) $\overline{K \cup L} = \overline{K} \cup \overline{L}$,
(iii) $\overline{K \cap L} \subseteq \overline{K} \cap \overline{L}$.

Proof. (i) Considering the regular substitution $\varphi : A^* \to A^*$ defined by the rule $\varphi(a) = a^+$ for any $a \in A$, we can write $\overline{L} = \varphi(\varphi^{-1}(L))$. Then we can apply Theorem 4.4. from [12] saying that the family of regular languages is closed under regular substitutions and inverse regular substitutions.

(ii) and (iii) are trivial observations. □

Also the following is obvious.

Lemma 2. *For a regular $L \subseteq A^*$, the following statements are equivalent :*
(i) L is literally idempotent,
(ii) $\overline{L} = L$,
(iii) $\sim\ \subseteq\ \sim_L$,
(iv) the (canonical) minimal DFA for L is literally idempotent,
(v) $(\ \forall\ u \in A^*,\ a \in A\ )\ (a^2)^{-1}(u^{-1}L) = a^{-1}(u^{-1}L)$,
(vi) L is a (disjoint) union (not necessarily finite !) of the languages of the form

$$a_1^+ a_2^+ \dots a_k^+,\ k \in \mathbb{N}_0,\ a_1, \dots, a_k \in A,\ a_1 \neq a_2 \neq \dots \neq a_k\ .$$

For a class of languages $\mathscr{C}$, we can consider the class of literally idempotent languages $\overline{\mathscr{C}}$ where $\overline{\mathscr{C}}(A) = \{ \ \overline{L} \mid L \in \mathscr{C}(A) \ \}$ for each A. Clearly, the following holds.

Lemma 3. *Let $\mathscr{C}$ be a class of languages. Then :*
(i) A class $\overline{\mathscr{C}}$ is closed under union whenever $\mathscr{C}$ is closed under union.
(ii) $\mathscr{C} \cap \mathscr{L} \subseteq \overline{\mathscr{C}}$.

5 Varieties of Literally Idempotent Languages

Our main results consist in syntactic characterizations of certain classes of languages – see Propositions 1, 2 and 4, together with the following result – see Propositions 1, 2, 3, 4, 5.

Theorem 1. *For the class $\mathscr{V} \in \{\mathscr{V}_{1/2}, \mathscr{V}^c_{1/2}, \mathscr{R}, \mathcal{V}_1, \mathscr{V}_{3/2}\}$ we have $\mathscr{V} \cap \mathscr{L} = \overline{\mathscr{V}}$.*

On the contrary, the following examples show that $\overline{\mathscr{V}}$ need not be a (positive) literal variety if $\mathscr{V}$ is a (positive) literal variety.

Example 1. We consider the class $\mathscr{N}_f$. Now $\mathscr{N}_f \cap \mathscr{L}$ consists of full languages, the empty language and the unit language, i.e. $(\mathscr{N}_f \cap \mathscr{L})(A) = \{\emptyset, \{\epsilon\}, A^*\}$. It is an easy observation that this literal variety is given by the literal pseudoidentity $x = y$.

On the other hand, $\emptyset, \{\epsilon\}, A^* \in \overline{\mathscr{N}_f}(A)$ and a language $L \notin \{\emptyset, A^*\}$ over A belongs to $\overline{\mathscr{N}_f}(A)$ iff L is a finite union of languages of the form $a_1^+ a_2^+ \dots a_n^+$ where $a_1, \dots, a_n \in A$, $a_1 \neq a_2 \neq \dots \neq a_n$ (for $n = 0$ we mean the language $\{\epsilon\}$). This implies that $\overline{\mathscr{N}_f}$ is not a literal positive variety of languages, because $\overline{\mathscr{N}_f}$ is not closed under inverse literal homomorphic images. If we consider the literal positive variety of languages generated by $\overline{\mathscr{N}_f}$ then it is easy to see that L belongs to $< \overline{\mathscr{N}_f} >_{plv} (A)$ if and only if L is a finite union of languages of the form $B_1^+ B_2^+ \dots B_n^+$ where $\{B_1, B_2, \dots, B_n\}$ is a partition of a subset of the alphabet A (i.e. different B_i's are disjoint) and $B_1 \neq \dots \neq B_n$. One can show that this positive variety is given by the literal satisfaction of the following pseudoidentities and pseudoinequalities

$$x^2 = x, \ u^\omega v x = u^\omega v y, \ x v u^\omega = y v u^\omega, \ x \leq u^\omega \ ,$$

$$\text{for all } u, v \in \Sigma^+ \text{ such that } x, y \in c(u), \ x, y \in \Sigma \ .$$

Example 2. We can also consider the variety $\mathscr{N}$. Now

$$(\mathscr{N} \cap \mathscr{L})(A) = \{ \, \emptyset, \{\epsilon\}, A^+, A^* \, \}$$

Moreover, if the language L over A is cofinite then $\overline{L} \in \{A^+, A^*\}$. From this reason $\overline{\mathscr{N}}(A) = \overline{\mathscr{N}_f}(A) \cup \{A^+\}$ and again it is not a literal variety.

Now, we will study the new classes from Section 3. We start with the variety $\mathscr{V}_{1/2}$. For a word $u = a_1 a_2 \dots a_k, \ a_1, \dots, a_k \in A$, we denote

$$L_u = A^* a_1 A^* a_2 \dots a_k A^*$$

– the set of all words which contain the word u as a subword.

Proposition 1. *For a language L over A, the following are equivalent :*

(i) L is a finite union of languages of the form $(L\ 1/2)$,

(ii) $L \in (\mathscr{V}_{1/2} \cap \mathscr{L})(A)$,

(iii) the syntactic homomorphism $\phi_L : A^* \to \mathsf{O}(L)$ of the language L satisfies the pseudoinequalities $x \leq 1$ and $x^2 = x$ literally,

(iv) $L \in \overline{\mathscr{V}_{1/2}}(A)$.

Proof. "(i) $\implies$ (ii)": If the language L is a finite union of languages of the form $(L\ 1/2)$ then it belongs to $\mathscr{V}_{1/2}(A)$ and we will check that it is also a literally idempotent language. Indeed, for $L = A^*a_1A^*a_2 \dots a_kA^*$ where $a_1 \neq \dots \neq a_k$, we have $a_1^{-1}L = L + A^*a_2A^*a_3 \dots a_kA^*$ and consequently $a_1^{-2}L = a_1^{-1}L$; for $a \in A$, $a \neq a_1$ we have $a^{-1}L = L$ and $a^{-2}L = a^{-1}L$ follows.

"(ii) $\Leftrightarrow$ (iii)" is clear because ϕ_L satisfies the pseudoinequality $x \leq 1$ literally iff $\mathsf{O}(L)$ satisfies this pseudoinequality in the classical sense.

"(ii) $\implies$ (iv)" follows from Lemma 3.

"(iv) $\implies$ (i)": If $L \in \overline{\mathscr{V}_{1/2}}(A)$ then L is a finite union of languages of the form $\overline{L}_u$. We prove that $\overline{L}_u$ is of the form $(L\ 1/2)$.

First, we claim that $\overline{L}_u = \overline{L}_{\overrightarrow{u}}$. The inclusion $L_u \subseteq L_{\overrightarrow{u}}$ is trivial and $\overline{L}_u \subseteq \overline{L}_{\overrightarrow{u}}$ follows. Now, assume that $w \in \overline{L}_{\overrightarrow{u}}$ then there is a word $s \in L_{\overrightarrow{u}}$ such that $w \sim s$. We define the word $s_{|u|}$ in such a way, that we replace any letter a in s by $a^{|u|}$, where $|u|$ is a length of the word u. Because s contains the word $\overrightarrow{u}$ as a subword, we can see that $s_{|u|}$ contains the word u. Hence $w \sim s_{|u|} \in L_u$ and we can conclude $w \in \overline{L}_u$.

We proved that $\overline{L}_u = \overline{L}_{\overrightarrow{u}}$ and because $L_{\overrightarrow{u}}$ is of the form $(L\ 1/2)$, i.e. it is literally idempotent as we proved at the beginning of the proof, we have also $\overline{L}_{\overrightarrow{u}} = L_{\overrightarrow{u}}$ which implies that $\overline{L}_u$ is of the form $(L\ 1/2)$. □

We prove now a similar theorem for the class $\mathscr{V}^c_{1/2}$ (here $\mathscr{V}^c(A) = \{\, A^* \setminus L \mid L \in \mathscr{V}(A) \,\}$ for each A). At first, we formulate the following technical lemma which describes basic properties of languages of the form L^c_u.

Lemma 4. *Let* $u, u_1, \dots, u_k, w \in A^*,\ k \in \mathbb{N},\ u = a_1a_2 \dots, a_k,\ a_1, \dots, a_k \in A.$ *Then :*

(i) $w \in \overline{L^c_u} \ \Leftrightarrow\ \overrightarrow{w} \in L^c_u$.

(ii) $\overline{L^c_u} = (A \setminus \{a_1\})^*\ a_1^*\ (A \setminus \{a_2\})^*\ a_2^*\ \dots\ a_{k-1}^*\ (A \setminus \{a_k\})^*$.

(iii) $\overline{L^c_{u_1} \cap \dots \cap L^c_{u_n}} = \overline{L^c_{u_1}} \cap \dots \cap \overline{L^c_{u_n}}$.

Proof. (i) The implication "$\Longleftarrow$" is trivial.

"$\Longrightarrow$" : If $w \in \overline{L^c_u}$ then there is a word $v \in L^c_u$ such that $w \sim v$. This means that v does not contain the word u as a subword. Hence $\overrightarrow{v}$ does not contain the word u as a subword too, i.e. $\overrightarrow{w} = \overrightarrow{v} \in L^c_u$.

(ii) We denote $K = (A \setminus \{a_1\})^*\ a_1^*\ (A \setminus \{a_2\})^*\ a_2^*\ \dots\ a_{k-1}^*\ (A \setminus \{a_k\})^*$ and we will prove that $\overline{L^c_u} = K$.

"$\subseteq$" : If $w \in \overline{L^c_u}$ then $\overrightarrow{w} \in L^c_u$ by (i). If we read $\overrightarrow{w}$ from left to right and look for the first occurrence of a_1 (if it exists) and then look for the first occurrence of

a_2 after this first occurrence of a_1 (if it exists) and so on, we obtain the following factorization of $\overrightarrow{w}$:

$$\overrightarrow{w} = w_1 a_1 w_2 a_2 \dots a_l w_{l+1}, \text{ where } l < k, w_i \in (A \setminus \{a_i\})^* .$$

Hence $\overrightarrow{w} \in K$ and because $\overline{K} = K$ we have also $w \in K$.

"$\supseteq$" : Let $w \in K$. Then $w = w_1 a_1^{\alpha_1} w_2 a_2^{\alpha_2} \dots a_{k-1}^{\alpha_{k-1}} w_k$ where $w_i \in (A \setminus \{a_i\})^*$ and we can also assume without the lost of generality that w_i does not start with the letter a_{i-1}. Hence $\overrightarrow{w} = \overrightarrow{w_1} a_1 \overrightarrow{w_2} a_2 \dots a_{k-1} \overrightarrow{w_k}$ and one can check by induction with respect to i that the word $\overrightarrow{w_1} a_1 \overrightarrow{w_2} a_2 \dots a_{i-1} \overrightarrow{w_i}$ does not contain the word $a_1 \dots a_i$ as a subword. This implies that $\overrightarrow{w} \in L_u^c$ and $w \in \overline{L_u^c}$ follows by (i).

(iii) The inclusion "$\subseteq$" is a trivial consequence of Lemma 1 (iii) and the inclusion "$\supseteq$" is a consequence of (i). Indeed, $w \in \overline{L_{u_1}^c} \cap \dots \cap \overline{L_{u_n}^c}$ implies $w \in \overline{L_{u_i}^c}$ and $\overrightarrow{w} \in L_{u_i}^c$, for $i = 1, \dots, n$, follows. Hence $\overrightarrow{w} \in L_{u_1}^c \cap \dots \cap L_{u_n}^c$ and consequently $w \in \overline{L_{u_1}^c \cap \dots \cap L_{u_n}^c}$. □

Proposition 2. *For a language L over A, the following are equivalent :*

(i) L is a finite union of the languages of the form $(L\ 1/2\ c)$.

(ii) $L \in (\mathscr{V}_{1/2}^c \cap \mathscr{L})(A)$.

(iii) The syntactic homomorphism $\phi_L : A^* \to \mathsf{O}(L)$ satisfies the pseudoinequalities $x^2 = x$ and $1 \le x$ literally.

(iv) $L \in \overline{\mathscr{V}_{1/2}^c}(A)$.

(v) $L \in \overline{\mathscr{V}_{1/2}}^c(A)$.

(vi) L is a finite intersection of the languages of the form $(L\ 1/2\ c)$.

Proof. "(i) $\implies$ (iii)": Let L be a finite union of the languages of the form $(L\ 1/2\ c)$. For any $K = B_1^* B_2^* \dots B_k^*$ and any $a \in A$ we have $a^{-1}K = B_i^* \dots B_k^*$ where i is the smallest index with the property $a \in B_i$. Now we see that $a^{-2}K = a^{-1}K$ and moreover $K \supseteq a^{-1}K$. ¿From this we can conclude $a^{-2}L = a^{-1}L$, $L \supseteq a^{-1}L$. The first observation implies L is literally idempotent and the second one implies that $a \geq 1$ holds in $\mathsf{O}(L)$ for any $a \in A$. In other words, the syntactic homomorphism $\phi_L : A^* \to \mathsf{O}(L)$ satisfies the pseudoinequalities $x^2 = x$ and $1 \le x$ literally.

As in the previous proof we have that (iii) is equivalent to (ii) and (ii) implies (iv) by Lemma 3.

"(iv) $\implies$ (vi)": Let $L \in \overline{\mathscr{V}_{1/2}^c}(A)$. Then $L = \overline{R}$, where $R \in \mathscr{V}_{1/2}^c$. So, R^c is a finite union of the languages of the form $A^* a_1 A^* a_2 \dots a_k A^*$, $k \in \mathbb{N}_0$, $a_1, \dots, a_k \in A$. This means that R is a finite intersection of the languages L_u^c. The language $L = \overline{R}$ is an intersection of languages of the form $\overline{L_u^c}$ by (iii) in Lemma 4. Moreover, any of these languages is of the form $(L\ 1/2\ c)$ by (ii) in the same lemma.

"(vi) $\implies$ (i)" is a consequence of the fact that any intersection of two languages of the form $(L\ 1/2\ c)$ is a finite union of the languages of the form $(L\ 1/2\ c)$.

So, we proved that the conditions (i) – (iv) and (vi) are equivalent. The condition (v) is equivalent to those by Proposition 1. □

Proposition 3. *For a language L over A, the following are equivalent :*

(i) L is a finite union of languages of the form $(L\ R)$.

(ii) $L \in (\mathscr{R} \cap \mathscr{L})(A)$.

(iii) $L \in \overline{\mathscr{R}}(A)$.

Proof. "(i) $\implies$ (ii)" is similar to the previous proofs.

"(ii) $\implies$ (iii)" : Again by Lemma 3.

"(iii) $\implies$ (i)" : If $L \in \overline{\mathscr{R}}(A)$, then $L = \overline{R}$, where $R \in \mathscr{R}$. So, R is a finite union of the languages of the form (R). We show that any such language $\overline{B_0^* a_1 B_1^* a_2 \dots a_k B_k^*}$ can be written as a finite union of languages of the form $(L\ R)$. We prove that any language of the form $\overline{B_0^* a_1 B_1^* a_2 \dots a_k B_k^*}$ with the set of "bad" indices $\{\, i \mid a_i = a_{i+1} \text{ or } a_i \notin B_i \,\}$ can be written as a union of languages of the same form, but with the smaller set of "bad" indices. Hence we can inductively rewrite any given language $\overline{B_0^* a_1 B_1^* a_2 \dots a_k B_k^*}$ as a finite union of languages of the form $(L\ R)$.

Let $K = \overline{B_0^* a_1 B_1^* a_2 \dots a_k B_k^*}$ be a language and i be such that $a_i = a_{i+1}$ or $a_i \notin B_i$.

First, assume that $a_i = a_{i+1}$. Then $a_i \notin B_i$. If $B_i = \emptyset$ we can simple remove $B_i^* a_{i+1}$ from the expression of the language K. Otherwise we write the language K as a union of certain languages L_c for $c \in B_i \cup \{a_i\}$ as follows. The language L_{a_i} comes from our expression if we exchange the string $a_i B_i^* a_{i+1}$ by a_i, i.e.

$$L_{a_i} = \overline{B_0^* a_1 B_1^* a_2 \dots a_{i-1} B_{i-1}^* a_i B_{i+1}^* \dots a_k B_k^*}\ .$$

This language consists of words from K which do not use letters from B_i. For $c \in B_i$ the language L_c comes from our expression if we exchange the part $a_i B_i^* a_{i+1}$ by $a_i a_i^* c B_i^* a_{i+1}$, i.e.

$$L_c = \overline{B_0^* a_1 B_1^* a_2 \dots a_{i-1} B_{i-1}^* a_i a_i^* c B_i^* a_{i+1} B_{i+1}^* \dots a_k B_k^*}.$$

The language L_c consists of words from K which use letters from B_i and the first such letter is c.

In the second case we have $a_i \neq a_{i+1}$ and $a_i \notin B_i$ and we can apply a similar construction. □

For a class $\mathscr{V}$ of languages we put :
$\mathscr{V}^d(A) = \{\, L^d \mid L \in \mathscr{V}(A) \,\}$ – the class *dual* to $\mathscr{V}$, where
$L^d = \{\, a_n \dots a_1 \mid a_1, \dots a_n \in L,\ a_1, \dots, a_n \in A \,\}$ – the language dual to L.

Corollary 1. $\mathscr{V}_1 \cap \mathscr{L} = \overline{\mathscr{V}_1}$.

Proof. We have $\mathscr{R} \cap \mathscr{L} = \overline{\mathscr{R}}$ which has the dual version, i.e. $\mathscr{R}^d \cap \mathscr{L} = \overline{\mathscr{R}^d}$. It is well known that $\mathscr{R} \cap \mathscr{R}^d = \mathscr{V}_1$.

The inclusion $\mathscr{V}_1 \cap \mathscr{L} \subseteq \overline{\mathscr{V}_1}$ follows from Lemma 3.

If $L \in \overline{\mathscr{V}_1}(A)$ then $L = \overline{L_1}$ where $L_1 \in \mathscr{V}_1(A)$. Hence $L_1 \in \mathscr{R}(A)$, and $L_1 \in \mathscr{R}^d(A)$ and we obtain $L = \overline{L}_1 \in \overline{\mathscr{R}}(A)$, $L = \overline{L}_1 \in \overline{\mathscr{R}^d}(A)$. Now we use previous theorem and obtain $L \in (\mathscr{R} \cap \mathscr{L})(A)$ and $L \in (\mathscr{R}^d \cap \mathscr{L})(A)$. Hence $L \in (\mathscr{R} \cap \mathscr{R}^d \cap \mathscr{L})(A) = (\mathscr{V}_1 \cap \mathscr{L})(A)$. □

Proposition 4. *For a language L over A, the following are equivalent :*

(i) L is a boolean combination of languages of the form (L 1/2).

(ii) L is a boolean combination of the languages of the form (L 1/2 c).

(iii) $L \in (\mathscr{V}_1 \cap \mathscr{L})(A)$.

(iv) The syntactic homomorphism $\phi_L : A^* \to \mathsf{O}\,(L)$ of the language L satisfies the pseudoidentity $x^2 = x$ literally and $\mathsf{O}\,(L)$ is $\mathcal{J}$-trivial.

(v) $L \in \overline{\mathscr{V}_1}(A)$.

Proof. The conditions (i) and (ii) are equivalent by Proposition 2. The equivalence of conditions (iii) and (iv) follows from the characterization of varieties $\mathscr{V}_1$ and $\mathscr{L}$. The equivalence of conditions (iii) and (v) is contained in Corollary 1. The implication (i) $\Longrightarrow$ (iii) holds as $\mathscr{V}_1 \cap \mathscr{L}$ is closed under boolean operations. In the rest we show the implication (iii) $\Longrightarrow$ (i) which concludes the proof.

Let $L \in (\mathscr{V}_1 \cap \mathscr{L})(A)$. Then L is a boolean combination of the languages of the form $A^*a_1A^*a_2 \dots a_kA^*,\ k \in \mathbb{N}_0,\ a_1, \dots, a_k \in A$ and moreover L is literally idempotent, i.e. $L = \overline{L}$ is a finite union of the languages of the form

$$\overline{L_{u_1} \cap \dots \cap L_{u_n} \cap L^c_{v_1} \cap \dots \cap L^c_{v_k}}.$$

We will show that this language can be written as a boolean combination of the languages of the form (L 1/2). In fact we will follow a decomposition of a language from the class $\mathscr{V}_1$ to a boolean combination. Because our literally idempotent language is fully given by the words in normal form contained in it, we will concentrate on such words.

Let $K = L_{u_1} \cap \dots \cap L_{u_n} \cap L^c_{v_1} \cap \dots \cap L^c_{v_k}$. We denote r the maximal length of words in the set $\{u_1, \dots, u_n, v_1, \dots v_k\}$. Now for any word $w \in \overline{K}$ in the normal form, and of the length at least $2r$, we consider two following lists of words in the normal forms :

$s_1, \dots, s_p$ are all words of the length $2r$, in the normal form which are subwords of w.

$t_1, \dots, t_q$ are all words of the length $2r$ in the normal form which are not contained in w as subwords.

We consider the language

$$N_w = \overline{L_{s_1}} \cap \dots \cap \overline{L_{s_p}} \cap \overline{L^c_{t_1}} \cap \dots \cap \overline{L^c_{t_q}}.$$

In this way we define finitely many languages (for all w's we have only finitely many s's and t's). By Proposition 2, all are boolean combinations of languages of the form (L 1/2). For a word $z \in \overline{K}$, in the normal form, and of the length smaller than $2r$ we simply denote $M_z = \overline{\{z\}}$. Note that $M_z = a_1^+a_2^+ \dots a_l^+$ where $z = a_1 \dots a_l$ and this language is of our form because $M_z = a_1^* \dots a_l^* \cap L_z$.

We will show that

$$\overline{K} = \bigcup_w N_w \ \cup \ \bigcup_z M_z \ .$$

"$\subseteq$" : Let $x \in \overline{K}$. Then $w = \overrightarrow{x} \in \overline{K}$. If w has the length smaller than $2r$ then $x \in M_w$. If w has the length at least $2r$ then $x \in N_w$.

"$\supseteq$" : If $x \in M_z$ then $x \sim z \in \overline{K}$ and $x \in \overline{K}$ follows.

If $x \in N_w$ then we have $x \in \overline{L_{s_i}}$, and $x \in \overline{L^c_{t_j}}$ for all i's and j's . We want to prove that $\overrightarrow{x} \in K$, i.e. $\overrightarrow{x} \in L_{u_i}$ and $\overrightarrow{x} \notin L_{v_j}$.

If we take an arbitrary $u \in \{u_1, \ldots, u_n\}$ then u is contained in w. Because w is in the normal form and the length of u is $\leq r$, we can find u as a subword of the word s, which is a subword of w, it is a word in the normal form and it has the length $2r$. This means that u is a subword of some s_i which is a subword of $\overrightarrow{x}$. Hence $\overrightarrow{x} \in L_u$

Now we take an arbitrary $v \in \{v_1, \ldots, v_k\}$ and assume for a moment that $\overrightarrow{x} \in L_v$. This means that $\overrightarrow{x}$ contains v as a subword. Again we can find a subword s of the word $\overrightarrow{x}$ such that s is in the normal form, contains v as a subword and it is of length $2r$. Because $\overrightarrow{x} \in N_w$ we know that $s \in \{s_1, \ldots, s_p\}$. Hence v is a subword of s, which is a subword of $w \in \overline{K}$. This implies that for any word y such that $\overrightarrow{y} = w$ we have $y \in L_v$ and consequently $y \in K$. This is a contradiction with $w \in \overline{K}$. $\square$

One can prove the following result in the similar way as Proposition 4.

Proposition 5. *For a language L over A, the following are equivalent :*
(i) L is a finite union of languages of the form $(L\ 3/2)$.
(ii) $L \in (\mathscr{V}_{3/2} \cap \mathscr{L})(A)$.
(iii) $L \in \overline{\mathscr{V}_{3/2}}(A)$.

Example 3. We can consider similar variety of all languages which are finite unions of languages of the form $(L\ E)$. It is clear that this class is a literal positive variety contained in $\overline{\mathscr{V}_{3/2}}$. The inclusion is proper as we have an example of the language $a^*b^+ = a^*bb^* \in \overline{\mathscr{V}_{3/2}}$ which can not be written in the previous way.

6 Literally Idempotent Languages over Two Letter Alphabet

If we consider one letter alphabet $\{a\}$, then the literally idempotent languages are exactly $\emptyset, \{\epsilon\}, a^+, a^*$.

It is well-known that for a regular languages L over the alphabet $\{a\}$ the set $\{\, i \mid a^i \in L \,\}$ is semilinear (i.e. it is a finite union of linear sets). In other words, the language L can be written in the form $L = A \cup B \cdot (a^k)^*$, where $k \in \mathbb{N}_0$ and A and B are finite languages. Moreover, such expression of the language L can be chosen in a canonical way. (All these observations can be easily established if one can look at the minimal automaton of the language L.)

We indicate that the similar canonical form can be given in the case of the literally idempotent languages over the two letter alphabet $\{a, b\}$.

Let L be an arbitral literally idempotent language over the alphabet $\{a, b\}$. This language is uniquely determined by the words in normal forms, hence it is natural to consider the sets:

$$_aI_a = \{\, i \in \mathbb{N}_0 \mid (ab)^i a \in L \,\},\ _aI_b = \{\, i \in \mathbb{N}_0 \mid (ab)^i \in L \,\},$$

$$_bI_a = \{\, i \in \mathbb{N} \mid (ba)^i \in L \,\},\ _bI_b = \{\, i \in \mathbb{N}_0 \mid (ba)^i b \in L \,\}\ .$$

If we look at the minimal automaton of the language L, which is literally idempotent, we see that any of the sets $_aI_a$, $_aI_b$, $_bI_a$ and $_bI_b$ is semilinear. This observations lead to expression of the language $L \cap a\{a,b\}^*a$ (and $L \cap a\{a,b\}^*b$, $L \cap b\{a,b\}^*a$, $L \cap b\{a,b\}^*b$ respectively) in the form $\overline{A} \cup \overline{B}((a^+b^+)^k)^*$, where $k \in \mathbb{N}_0$ and A and B are finite languages over the alphabet $\{a,b\}$. Moreover such expression of the language $L \cap a\{a,b\}^*a$ can be chosen in a canonical way, if we add assumptions that A, B are the smallest possible ones. Note that $L \cap a\{a,b\}^*a$ is finite if and only if B is empty.

7 Linear Temporal Logic without next

In this section we introduce a connection between the Linear Temporal Logic (LTL) and the concept of literal idempotency. The expressive power of certain variants of the temporal logics were successfully characterized applying algebraic methods, in particular the concept of the syntactic monoid, in [2], [3], and [11]. In the center of our interest is the expressive power of LTL formulas which do not use the "next" operator.

First, we recall basic definitions. *A formula of linear temporal logic without next operator* (LTLWN) over a finite set A of letters is built from the elements of the alphabet A and the logical constant $\mathbb{T}$ (true) using the boolean connectives $\neg$ and $\vee$ and the temporal logic operator $\mathbb{U}$ (until).

Let $w \in A^*$ be a word over A. The length of w is denoted by $|w|$. For any $1 \le i \le n = |w|$ we denote by $w(i)$ the i-th letter of w and w_i the suffix of w starting at the i-th position, i.e. $w_i = w(i)w(i+1)\dots w(n)$.

The validity of the formula φ of LTLWN on $w \in A^*$ is defined as follows :
$w \models \mathbb{T}$
$w \models a \ \Leftrightarrow\ w(1) = a$
$w \models \neg\varphi \ \Leftrightarrow\ w \not\models \varphi$
$w \models \varphi_1 \vee \varphi_2 \ \Leftrightarrow\ w \models \varphi_1 \vee w \models \varphi_2$
$w \models \varphi_1 \mathbb{U} \varphi_2 \ \Leftrightarrow\ \exists i \in \mathbb{N} : w_i \models \varphi_2 \wedge \forall 1 \le j < i : w_j \models \varphi_1$.

Every formula φ defines the language $L_\varphi = \{\, w \in A^* \mid w \models \varphi \,\}$.

It is well-known that language is definable by linear logic formula iff it is star-free, i.e. it has aperiodic syntactic monoid.

The first easy observation, which is mentioned in different way in literature, e.g. in [6], is the following statement.

Result 3 *If a language L is definable by a formula of the Linear Temporal Logic without next operator, i.e. $L \in LTLWN$, then it is literally idempotent.*

The class of languages which are definable without until and next operators is well-known. The interesting point of view is, that this class, denoted by U_0, forms a literal variety, which is characterized in the following lemma.

Lemma 5. *Let L be a language. Then $L \in U_0$ if and only if $\varphi_L \models_{\mathfrak{L}} xy = x$.*

Proof. It is not hard to see that $L \in U_0$ iff it is of the form BA^*, or $BA^* \cup \{\epsilon\}$, where $B \subseteq A$. Indeed, For an non-empty subset $B = \{b_1, \dots, b_k\} \subseteq A$ we have the formula $\varphi_B = b_1 \vee b_2 \vee \dots \vee b_k$ which defines the language BA^*, and also the formula $\varphi_\emptyset = \neg\mathbb{T}$, which defines the language BA^* for $B = \emptyset$. We can also put $\varphi_\epsilon = \neg\varphi_A$, a formula getting the language $\{\epsilon\}$.

Now if we take an arbitral language of the form BA^*, or $BA^* \cup \{\epsilon\}$ then the syntactic monoid of L has (at most) three elements $1, b, c$, with multiplication given by rules $bc = bb = b$, $cc = cb = c$. The letters from B are mapped by syntactic morphism to b and the letters from $A \setminus B$ are mapped to c. (If $B = \emptyset$ or $B = A$ then the syntactic monoid has at most two elements.) So, we see that the identity is satisfied. Also the opposite implication is easy to get. □

Note that in the previous lemma we can not exchange the literal validity of the identity $xy = x$ with the classical one, because in a classical sense the identity has a consequence $1 \cdot y = 1$. The reason is that the class U_0 is not closed under all homomorphic preimages, e.g. $\varphi^{-1}(\epsilon) = B^*$ for a morphism which maps letters from a subset $B \subseteq A$ to the empty word.

We add one more example of formulas of a special form which correspond to certain literal variety of languages.

We say that the formula φ is *easy* if it is of the form

$$\psi = \varphi_{B_1}\mathbb{U}(\varphi_{B_2}\mathbb{U}(\dots(\varphi_{B_n}\mathbb{U}\varphi_\epsilon))\dots).$$

The language is *easy* if it is definable by boolean combinations of easy formulas.

Lemma 6. *Let L be a language over A. Then L is easy if and only if $L \in \overline{\mathscr{V}_1}(A)$.*

Proof. It is easy to see that an easy formula $\psi = \varphi_{B_1}\mathbb{U}(\varphi_{B_2}\mathbb{U}(\dots(\varphi_{B_n}\mathbb{U}\varphi_\epsilon))\dots)$ define the language $B_1^*B_2^*\dots B_n^*$. Hence the statement follows from Proposition 4. □

Recall, that the syntactic monoid of the language $L = a^*b^*$ is not idempotent (as $ab \in L$ and $abab \notin L$) but the syntactic morphism satisfies idempotency literally. This could be a good motivation for a future work in the field of applications of literal idempotent varieties in LTL. The investigations of connections of certain hierarchies of languages with LTL are in progress.

References

1. J. Almeida, *Finite Semigroups and Universal Algebra*, World Scientific, 1994
2. J. Cohen, J.–E. Pin and D. Perrin, On the expressive power of temporal logic, *J. Computer and System Science* 46 (1993), 271–294.
3. Z. Ésik, Extended temporal logic on finite words and wreath product of monoids with distinguished generators, *Proc. DLT 02*, LNCS 2450, Springer, 2003, 43–58.
4. Z. Ésik and M. Ito, Temporal logic with cyclic counting and the degree of aperiodicity of finite automata, *Acta Cybernetica* **16** (2003), 1–28, a preprint BRICS 2001

5. Z. Ésik and K.G. Larsen, Regular languages defined by Lindström quantifiers, *Theoretical Informatics and Applications* **37** (2003), 197–242, preprint BRICS 2002
6. A. Kučera and J. Strejček, The stuttering principle revisited, *Acta Informatica*, **41/7** (2005) 415–434
7. M. Kunc, Equational description of pseudovarieties of homomorphisms, *Theoretical Informatics and Applications* **37** (2003), 243–254
8. J.-E. Pin, *Varieties of Formal Languages*, Plenum, 1986
9. J.-E. Pin, Syntactic semigroups, Chapter 10 in *Handbook of Formal Languages*, G. Rozenberg and A. Salomaa eds, Springer, 1997
10. H. Straubing, On logical descriptions of regular languages, *Proc. Latin 2002*, Springer Lecture Notes in Computer Science, Vol. 2286, 2002, 528–538
11. D. Thérien and T. Wilke, Nesting until and since in linear temporal logic, *Theory of Computing Systems*, 37(1) 2003, 111-131.
12. S. Yu, Regular languages, Chapter 2 in *Handbook of Formal Languages*, G. Rozenberg and A. Salomaa eds, Springer, 1997