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/~polak

*Supported by the Grant Agency of the Czech Republic

A language $L \subseteq A^*$ is **literally idempotent** in case that $ua^2v \in L$ if and only if $uav \in L$ for each $u, v \in A^*$, $a \in A$.

A language $L \subseteq A^*$ is **literally idempotent** in case that $ua^2v \in L$ if and only if $uav \in L$ for each $u, v \in A^*$, $a \in A$.

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.

A language $L \subseteq A^*$ is **literally idempotent** in case that $ua^2v \in L$ if and only if $uav \in L$ for each $u, v \in A^*$, $a \in A$.

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.

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

$f : B^* \to A^*,\ K \in \mathscr{V}(A)$ 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.

$f : B^* \to A^*,\ K \in \mathscr{V}(A)$ 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
$f : B^* \to A^*$ 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.

$f : B^* \to A^*,\ K \in \mathscr{V}(A)$ 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 $f : B^* \to A^*$ 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 $\mathcal{M}$ (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 or Straubing for a more general notion of a $\mathbb{C}$-pseudovariety. Similarly, we define the literal pseudovarieties in the ordered case.

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

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

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 (Ésik and Larsen, Straubing) 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.

varieties of languages and literal pseudovarieties of homomorphisms onto finite monoids.

Let I_n, for $n \in \mathbb{N}$, be the set of all n-ary implicit operations for the class of finite monoids. 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

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

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

Result (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.

(i) Finite unions of languages

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

(i) Finite unions of languages

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

(ii) Finite unions of languages

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

(i) Finite unions of languages

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

(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

$$B_0^* a_1 B_1^* a_2 \dots a_k B_k^*,\ k \in \mathbb{N}_0,\ a_1, \dots, a_k \in A,\ B_0, \dots, B_k \subseteq A,$$

$$a_1 \neq a_2 \neq \dots \neq a_k,\ a_1 \in B_1, \dots, a_k \in B_k\ . \qquad (L\ 3/2)$$

(iv) Finite unions of languages of the form

$$B_0^* a_1 B_1^* a_2 \ldots a_k B_k^*,\ k \in \mathbb{N}_0,\ a_1, \ldots, a_k \in A,\ B_0, \ldots, B_k \subseteq A,$$

$$a_1 \neq a_2 \neq \cdots \neq a_k,\ a_1 \in B_1, \ldots, a_k \in B_k\ . \qquad (L\ 3/2)$$

(v) Finite unions of languages of the form

$$B_0^* a_1 B_1^* a_2 \ldots a_k B_k^*,\ k \in \mathbb{N}_0,\ a_1, \ldots, a_k \in A,\ B_0, \ldots, B_k \subseteq A,$$

$$a_1 \neq a_2 \neq \cdots \neq a_k,\ B_0 \not\ni a_1 \in B_1 \not\ni a_2 \in \cdots \not\ni a_k \in B_k\ . \quad (L\ R)$$

(vi) Finite unions of languages of the form

$$B_0^* a_1 B_1^* a_2 \dots a_k B_k^*,\ k \in \mathbb{N}_0,\ a_1, \dots, a_k \in A,\ B_0, \dots, B_k \subseteq A$$

$$a_1 \neq a_2 \neq \dots \neq a_k,\ a_1 \in B_0 \cap B_1, \dots, a_k \in B_{k-1} \cap B_k\ . \quad (L\ E)$$

(vi) Finite unions of languages of the form

$$B_0^* a_1 B_1^* a_2 \dots a_k B_k^*, \ k \in \mathbb{N}_0, \ a_1, \dots, a_k \in A, \ B_0, \dots, B_k \subseteq A$$

$$a_1 \neq a_2 \neq \dots \neq a_k, \ a_1 \in B_0 \cap B_1, \dots, a_k \in B_{k-1} \cap B_k \ . \quad (L\ E)$$

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^*.

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^*.

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 intersection $\mathscr{C} \cap \mathscr{L}$.

The second possibility is to consider the following (closure) operator on languages. For any language $L \subseteq A^*$, we define

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

We extend naturally this operator to classes of languages.

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

Moreover, those languages are exactly the classes (i) – (v) described above.