First order formulas with modular predicates

Laura Chaubard
LIAFA, Université Paris VII and CNRS, Case 7014,
2 Place Jussieu, 75251 Paris Cedex 05, France.
Laura.Chaubard@liafa.jussieu.fr

Jean-Éric Pin
LIAFA, Université Paris VII and CNRS, Case 7014,
2 Place Jussieu, 75251 Paris Cedex 05, France.
Jean-Eric.Pin@liafa.jussieu.fr

Howard Straubing
Department of Computer Science, Boston College, Chestnut Hill, MA 02467, USA
straubin@cs.bc.edu

Abstract

Two results by Schützenberger (1965) and by McNaughton and Papert (1971) lead to a precise description of the expressive power of first order logic on words interpreted as ordered colored structures. In this paper, we study the expressive power of existential formulas and of Boolean combinations of existential formulas in a logic enriched by modular numerical predicates. We first give a combinatorial description of the corresponding regular languages, and then give an algebraic characterization in terms of their syntactic morphisms. It follows that one can effectively decide whether a given regular language is captured by one of these two fragments of first order logic. The proofs rely on nontrivial techniques of semigroup theory: stamps, derived categories and wreath products.

1. Introduction

There is by now an extensive literature on the expressive power of various fragments of first order logic interpreted on finite words. There are also known connections with several areas in mathematics and computer science, including finite semigroups, automata, descriptive set theory, complexity, circuits and communication complexity. Further, this research is a necessary step towards the study of richer structures like infinite words, trees or graphs. This paper is a contribution to this theory.

Let us briefly describe the framework of our results. We associate to each nonempty word $u = a_0a_1 \dots a_{|u|-1}$ over the alphabet A a relational structure

$$\mathfrak{M}_u = \{(0, 1, \dots, |u| - 1), <, (\mathbf{a})_{a\in A}\}$$

where $<$ is the usual order on the domain and $\mathbf{a}$ is a predicate giving the positions i such that $a_i = a$. For instance, if $u = abbaaba$, then $\mathbf{a} = \{0, 3, 4, 6\}$ and $\mathbf{b} = \{1, 2, 5\}$. Given a formula φ, the language defined by φ is $L(\varphi) = \{u \in A^+ \mid \mathfrak{M}_u \text{ satisfies } \varphi\}$. Since languages may contain the empty word, we make the convention that a language L of A^* is defined by φ if $L(\varphi) = L \cap A^+$.

McNaughton and Papert [11] showed that a language is first-order definable (in the signature $\{<, (\mathbf{a})_{a\in A}\}$) if and only if it is star-free. The decidability of this class of regular languages, denoted by $\mathbf{FO}[<]$, follows from a celebrated result of Schützenberger [20]: a regular language is *star-free* if and only if its syntactic monoid is *aperiodic*. Thomas [27] (see also [13]) refined this correspondence between first order logic and star-free languages by showing that the concatenation hierarchy of star-free languages is, level by level, in correspondence with the Σ_n-hierarchy of first order formulas. However, little is known about the decidability of these classes. It is not very difficult to decide whether or not a given regular language belongs to $\mathbf{\Sigma_1}[<]$. The decidability of the Boolean closure of this class, denoted by $\mathcal{B}\mathbf{\Sigma_1}[<]$, relies on a nontrivial algebraic result of Simon [23]. The decidability of $\mathbf{\Sigma_2}[<]$ was also proved by algebraic methods [1, 17], but the decidability of the upper levels $\mathcal{B}\mathbf{\Sigma_2}[<]$, $\mathbf{\Sigma_3}[<]$ and beyond is a major open problem.

Several enrichments to the vocabulary $<$ were considered in the literature. Let $k \geq 0$. Recall that a k-ary *numerical predicate symbol* associates to each $n \geq 0$ a subset of

$\{0, \ldots, n-1\}^k$. We view $(i_1, \ldots, i_k) \in \{0, \ldots, n-1\}^k$ as a word $\delta_0 \cdots \delta_{n-1}$ over the alphabet $\Delta = 2^{\{1,\ldots,k\}}$ by setting $\delta_j = \{r \mid i_r = j\}$. Thus each numerical predicate symbol gives rise to a language in Δ^*. We say the numerical predicate symbol is *regular* if the corresponding language is regular. (Note that if $k = 0$, $\{0, \ldots, n-1\}^k$ is the one-element set $\{\emptyset\}$.)

Let $0 < d$ and $r \in \mathbb{Z}/d\mathbb{Z}$. We define two numerical predicate symbols (the *modular predicates*): The unary symbol MOD_r^d assigns to n the set $\{i < n \mid i \bmod d = r\}$, and the 0-ary symbol D_r^d assigns $\{\emptyset\}$ to n if $n \bmod d = r$, and $\emptyset$ otherwise. The associated languages are $(\emptyset^d)^*\emptyset^{r-1}\{1\}\emptyset^*$ and $(\emptyset^d)^*\emptyset^r$, respectively, so these are regular numerical predicates. Equivalently, we could introduce a constant symbol m denoting the last position in a string, in which case D_r^d is equivalent to $\text{MOD}_{r-1}^d m$. (This is the notation that we shall adopt below.)

We denote by $\mathbf{FO}[< + \text{MOD}]$ the logic obtained by adjoining all modular predicates. This signature was considered implicitly in automata theory and explicitly in a recent paper by Ésik and Ito [6]. It should not be confused with first order logic with modular quantifiers.

The logic $\mathbf{FO}[<+\text{REG}]$ is obtained by adjoining all regular numerical predicate symbols. This logic was considered in [2, 10, 12, 25] in connection with circuit complexity.

It is not difficult to see that $\mathbf{FO}[< + \text{MOD}] = \mathbf{FO}[< + \text{REG}]$. However, the lower levels of the Σ_n-hierarchy differ for the three signatures. The decidability of $\mathbf{\Sigma_1}[< + \text{REG}]$ and $\mathcal{B}\mathbf{\Sigma_1}[< + \text{REG}]$ was established in [10]. In this paper, we establish the decidability of the fragments $\mathbf{\Sigma_1}[<+\text{MOD}]$ and $\mathcal{B}\mathbf{\Sigma_1}[< + \text{MOD}]$, a problem left open in [6]. The situation is summarized in the table below:

	$<$	$< + \text{MOD}$	$< + \text{REG}$
$\mathbf{\Sigma_1}$	DECIDABLE [13, 27]	DECIDABLE **New result**	DECIDABLE [8, 10, 21]
$\mathcal{B}\mathbf{\Sigma_1}$	DECIDABLE [23, 27]	DECIDABLE **New result**	DECIDABLE [10]
$\vdots$			
FO	DECIDABLE [11, 20]	DECIDABLE [2, 25]	DECIDABLE [2, 25]

Our paper is organized as follows. Section 2 presents the necessary background to understand our proofs. Our main decidability results on fragments of first order logic are proved in Section 3 for $\mathbf{\Sigma_1}[< + \text{MOD}]$ and in Section 4 for $\mathcal{B}\mathbf{\Sigma_1}[< + \text{MOD}]$. In the last section, we summarize our results and compare them with other decidability results.

2. The algebraic approach

In this section, we survey the algebraic approach to automata theory that is needed to state our main results. We briefly present Eilenberg's variety theory [4], its extension to the ordered case [15] and its more recent generalization to stamps [5, 6, 7, 16, 26], in a form suitable to our purpose.

2.1 Semigroups, monoids and stamps

A *semigroup* is a set equipped with a binary associative operation, denoted multiplicatively, or additively when the semigroup is commutative. A *monoid* is a semigroup with a unit element. An element e of a semigroup is *idempotent* if $e^2 = e$. In a finite semigroup, every element x has a unique idempotent power, denoted by x^ω.

An element s of a semigroup S is said to be *regular* if and only if there exists an element $\bar{s}$ of S, called an *inverse* of s such that $s\bar{s}s = s$ and $\bar{s}s\bar{s} = \bar{s}$.

Given two monoids M and N, a *monoid morphism* is a map $\varphi : M \to N$ satisfying $\varphi(1) = 1$ and $\varphi(uv) = \varphi(u)\varphi(v)$ for all u, v in M. A monoid M is a *submonoid* of a monoid N if there exists an injective morphism from M into N. A monoid N is a *quotient* of a monoid M if there exists a surjective morphism from M onto N. A monoid M *divides* a monoid N if M is a quotient of a submonoid of N. The *product* of two monoids M_1 and M_2 is the set $M_1 \times M_2$ equipped with the product $(x_1, x_2)(y_1, y_2) = (x_1y_1, x_2y_2)$.

An *ordered* semigroup is a semigroup equipped with a partial order compatible with the operation of the semigroup. An *order ideal* I of an ordered semigroup $(S, \leq)$ is a subset of S such that if $x \in I$ and $y \leq x$ then $y \in I$.

Morphisms of ordered semigroups are order-preserving morphisms of semigroups. The notions of *ordered subsemigroup*, quotient and product are readily adapted from their unordered version and easily extended to the monoid case.

A *relational morphism* between two monoids M and N is a relation $\tau \colon M \to N$ which satisfies

(1) for every $s \in M$, $\tau(s) \neq \emptyset$,

(2) for every $s_1, s_2 \in M$, $\tau(s_1)\tau(s_2) \subseteq \tau(s_1s_2)$,

(3) $1 \in \tau(1)$.

A *stamp* is a morphism from a finitely generated free monoid onto a finite monoid. A stamp $\varphi : A^* \to M$ is said to be *trivial* if M is the trivial monoid. An *ordered stamp* is a stamp onto an ordered monoid.

Let $\varphi : A^* \to M$ be a stamp and let $Z = \varphi(A)$. Then Z is an element of the monoid $\mathcal{P}(M)$ of subsets of M, equipped with the product $XY = \{xy \mid x \in X, y \in Y\}$. Since $\mathcal{P}(M)$ is finite, Z has an idempotent power. This justifies the following definition: the *stability index* of a stamp $\varphi : A^* \to M$ is the least positive integer such that $\varphi(A^s) = \varphi(A^{2s})$. The set $\varphi(A^s)$ is a subsemigroup of M called the *stable semigroup of* φ and the monoid $\varphi(A^s) \cup \{1\}$ is called the *stable monoid of* φ.

2.2 Stamps and languages

Stamps and ordered stamps can be seen as language recognizers in the following way. Let $\varphi : A^* \to M$ be a stamp. A language L over A^* is *recognized by* the stamp φ if there exists a subset F of M such that $L = \varphi^{-1}(F)$. If M is ordered, we require F to be an order ideal of M. By extension, we say that the (ordered) monoid M recognizes L if there exists a stamp $\varphi : A^* \to M$ recognizing L.

A language is said to be *recognizable* if it is recognized by some finite monoid. Kleene's theorem asserts that recognizable and regular languages coincide.

Given a language L over A^*, we define the *syntactic congruence* $\sim_L$ and the *syntactic preorder* $\leq_L$ as follows:

(1) $u \sim_L v$ iff for all $x, y \in A^*$, $xvy \in L \Leftrightarrow xuy \in L$,

(2) $u \leq_L v$ iff for all $x, y \in A^*$, $xvy \in L \Rightarrow xuy \in L$.

The monoid $A^*/{\sim_L}$ is the *syntactic monoid* of L and is denoted by $M(L)$. It can be ordered with the partial order relation induced by $\leq_L$, to form the *ordered syntactic monoid* of L. The natural morphism $\eta_L : A^* \to M(L)$ is called the *syntactic (ordered) stamp* of L. The syntactic monoid of L is the smallest monoid (with respect to the division order on monoids) that recognizes L. In particular, a language is regular if and only if its syntactic monoid is finite.

From now on, all semigroups and monoids will be either finite or free.

2.3 The variety approach

The general idea of the variety theory is to classify regular languages through the algebraic properties of their syntactic invariants. For this purpose, Eilenberg originally considered classes of finite monoids defined by equations, called *varieties*. This gave an appealing framework in which to study classes of recognizable languages closed under Boolean operations, quotients, and inverse morphisms.

However, our classes $\mathbf{\Sigma_1}[<\!+\mathrm{MOD}]$ and $\mathcal{B}\mathbf{\Sigma_1}[<\!+\mathrm{MOD}]$ are not closed under inverse morphisms and the first one is not even closed under complement. Still, they are closed under inverses of *length-multiplying* morphisms and it is possible to adapt Eilenberg's variety theory to this weaker setting. The price to pay is the shift from the syntactic monoid to the syntactic stamp (for $\mathcal{B}\mathbf{\Sigma_1}[<\!+\mathrm{MOD}]$) or to the syntactic ordered stamp (for $\mathbf{\Sigma_1}[< + \mathrm{MOD}]$). The general framework for this study is the theory of $\mathcal{C}$-varieties, recently introduced by Straubing [26].

We first recall the classical notion of varieties. A *variety of finite monoids* is a class of (finite) monoids closed under division and finite product. Varieties of finite semigroups and of finite *ordered* monoids are defined analogously.

We now turn to varieties of stamps. Recall that a morphism $f : A^* \to B^*$ is *length-multiplying* (lm for short) if there exists an integer k such that the image of each letter of A is a word of B^k. A stamp $\varphi : A^* \to M$ *lm-divides* a stamp $\psi : B^* \to N$ if there is a pair (f, η) (called an lm-division), where $f : A^* \to B^*$ is an lm-morphism, $\eta : N \to M$ is a partial surjective monoid morphism, and $\varphi = \eta \circ \psi \circ f$. If f is the identity on A^*, the pair (f, η) is simply called a division. If φ and ψ are ordered stamps, that is, if M and N are ordered monoids, η is required to be order-preserving.

$$\begin{tikzcd}
A^* \arrow[rr, "f"] \arrow[d, "\varphi"'] & & B^* \arrow[d, "\psi"] \\
M & \mathrm{Im}(\psi \circ f) \subseteq \arrow[l, "\eta"'] & N
\end{tikzcd}$$

Figure 1. A division diagram.

The *product* of two stamps $\varphi_1 : A^* \to M_1$ and $\varphi_2 : A^* \to M_2$ is the stamp φ with domain A^* defined by $\varphi(a) = (\varphi_1(a), \varphi_2(a))$. The range of φ is a submonoid of $M_1 \times M_2$.

An *lm-variety of stamps* is a class of stamps containing the trivial stamps and closed under lm-division and finite products. The definition of a variety of ordered stamps is similar. Note that if $\mathbf{V}$ is a variety of finite (ordered) monoids, then the class of all (ordered) stamps whose range is in $\mathbf{V}$ forms an lm-variety of (ordered) stamps, also denoted by $\mathbf{V}$.

We now come to the definition of varieties of languages. A *positive Boolean algebra* is a set of languages that is closed under finite union and finite intersection. If it is also closed under complement, it is called a *Boolean algebra*. Given a language L and a word u, we set

$$u^{-1}L = \{v \in A^* \mid uv \in L\}$$
$$Lu^{-1} = \{v \in A^* \mid vu \in L\}$$

A *class of recognizable languages* $\mathcal{V}$ assigns to each finite alphabet A a set $\mathcal{V}(A^*)$ of recognizable languages of A^*. A *positive variety of languages* is a class of recognizable languages $\mathcal{V}$ such that for any alphabets A and B,

(1) $\mathcal{V}(A^*)$ is a positive Boolean algebra,

(2) if $L \in \mathcal{V}(A^*)$ and $a \in A$ then $a^{-1}L, La^{-1} \in \mathcal{V}(A^*)$,

(3) if $\varphi\colon A^* \to B^*$ is a morphism, $L \in \mathcal{V}(B^*)$ implies $\varphi^{-1}(L) \in \mathcal{V}(A^*)$.

A *variety of languages* is a positive variety $\mathcal{V}$ such that, for each alphabet A, $\mathcal{V}(A^*)$ is closed under complement.

Positive lm-varieties and *lm-varieties* of languages are defined in the same way by weakening Condition (3) to

(3′) if $\varphi\colon A^* \to B^*$ is an lm-morphism, $L \in \mathcal{V}(B^*)$ implies $\varphi^{-1}(L) \in \mathcal{V}(A^*)$.

Given a variety of finite monoids $\mathbf{V}$, the class $\mathcal{V}$ of all languages recognized by a monoid in $\mathbf{V}$ is a variety of languages. Eilenberg's theorem [4] asserts that the correspondence $\mathbf{V} \to \mathcal{V}$ is one-to-one and onto.

Similarly, if $\mathbf{V}$ is a variety of finite ordered monoids, the class $\mathcal{V}$ of all languages recognized by an ordered monoid in $\mathbf{V}$ is a positive variety of languages. It is proved in [15] that the correspondence $\mathbf{V} \to \mathcal{V}$ is one-to-one and onto.

Finally, given an lm-variety of (ordered) stamps $\mathbf{V}$, the class $\mathcal{V}$ of all languages recognized by a stamp in $\mathbf{V}$ is a (positive) lm-variety of languages. It is proved in [26] that the correspondence $\mathbf{V} \to \mathcal{V}$ is one-to-one and onto.

2.4 Examples

Example 2.1 The trivial variety of monoids $\mathbf{I}$ consists only of one monoid, the trivial monoid. The corresponding variety of languages $\mathcal{I}$ is defined, for every alphabet A, by $\mathcal{I}(A^*) = \{\emptyset, A^*\}$.

Example 2.2 A semigroup S is *locally trivial* if $eSe = \{e\}$ for each idempotent e of S. The class of *locally trivial* semigroups form a variety of semigroups, denoted by $\mathbf{LI}$.

Example 2.3 Let us denote by $\mathbf{J}^+$ the class of all finite ordered monoids $(M, \leq)$ such that, for all $x \in M$, $x \leq 1$. One can show that $\mathbf{J}^+$ is a variety of ordered monoids and that a language belongs to $\mathcal{J}^+(A^*)$ if and only if it is a finite union of languages of the form $A^*a_1A^* \cdots a_kA^*$, where $k \geq 0$ and $a_1, \ldots, a_k$ are letters of A. Further, it is shown in [13] that $\mathcal{J}^+$ is equal to the class $\mathbf{\Sigma_1}[<]$.

Example 2.4 A monoid M is *$\mathcal{J}$-trivial* if division is a partial order on M, that is, if the conditions $uxv = y$ and $syt = x$ imply $x = y$. The class of *$\mathcal{J}$-trivial* monoids form a variety, denoted by $\mathbf{J}$. Simon's theorem [22] states that $\mathcal{J}(A^*)$ is the Boolean algebra generated by the languages of the form $A^*a_1A^* \cdots a_kA^*$, where $k \geq 0$ and $a_1, \ldots, a_k$ are letters of A. It follows from [27] that $\mathcal{J}$ is also equal to the class $\mathcal{B}\mathbf{\Sigma_1}[<]$.

Example 2.5 A monoid M is *aperiodic* if there exists an integer n such that, for every $x \in M$, $x^n = x^{n+1}$. The class of aperiodic monoids form a variety denoted by $\mathbf{A}$. The results of Schützenberger [20] and McNaughton and Papert [11] show that the corresponding variety of languages is the class of star-free languages, or in logical terms, the class $\mathbf{FO}[<]$.

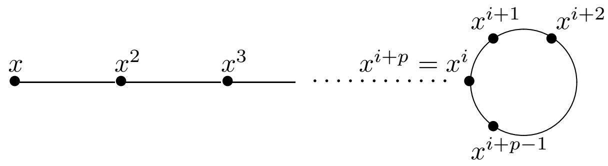
Example 2.6 Let $\mathbf{MOD}$ be the class of all stamps $\varphi : A^* \to M$ such that M is a cyclic group and $\varphi(a) = \varphi(b)$ for all letters a, b in A. Then $\mathbf{MOD}$ is an lm-variety of stamps. For each alphabet A, a language of $\mathcal{M}od(A^*)$ is recognized by some stamp $\pi_n : A^* \to \mathbb{Z}/n\mathbb{Z}$ and hence is a finite union of languages of the form $(A^n)^*A^k$ with $0 \leq k < n$.

Example 2.7 Given a variety of finite semigroups $\mathbf{V}$, a stamp is said to be a *quasi-$\mathbf{V}$ stamp* if its stable subsemigroup belongs to $\mathbf{V}$. It is stated in [26] that the quasi-$\mathbf{V}$ stamps form an lm-variety, denoted by $\mathbf{QV}$. It was proved in [2] that $\mathbf{FO}[< + \text{MOD}]$ is the lm-variety of languages corresponding to $\mathbf{QA}$.

2.5 Identities

Both varieties of finite monoids and lm-varieties of stamps have equational characterizations [19, 9, 16]. The same result holds for their ordered counterparts. The formal definition of identities requires the introduction of profinite topologies. Here we consider a simpler notion, illustrated with a few basic examples, which implies the result.

We start by recalling an elementary fact about finite semigroups. Let x be an element of a finite semigroup S. Since S is finite, there exist integers $i, p > 0$ such that $x^{i+p} = x^i$. The subsemigroup of S generated by x is represented below.



It is easy to see that the semigroup $\{x^i, \ldots, x^{i+p-1}\}$ is a cyclic group $G(x)$, whose identity is x^ω, the unique idempotent power of x.

An *ω-term* on an alphabet A is built from the letters of A using the usual concatenation product and two unary operators: $x \to x^\omega$ and $x \to x^{\omega-1}$. Thus, if $A = \{a, b, c\}$, abc, a^ω and $((ab^{\omega-1}c)^\omega ab)^\omega$ are examples of ω-terms.

Let $\varphi : A^* \to M$ be a stamp. The image $\varphi(t)$ of an ω-term t is defined recursively as follows. If t is a letter, then $\varphi(t)$ is already defined. If t and t' are ω-terms, then $\varphi(tt') = \varphi(t)\varphi(t')$. If $t = u^\omega$, then $\varphi(t)$ is the unique idempotent power of $\varphi(u)$. Finally if $t = u^{\omega-1}$, then $\varphi(t)$ is the inverse of $\varphi(u)^\omega\varphi(u)$ in the cyclic group $G(\varphi(u))$.

Let u, v be two ω-terms on a finite alphabet B. A stamp $\varphi : A^* \to M$ is said to *satisfy the lm-identity* $u = v$ if, for every lm-morphism $f\colon B^* \to A^*$, $\varphi \circ f(u) = \varphi \circ f(v)$. If M is ordered, we say that φ satisfies the lm-identity $u \leq v$ if, for every lm-morphism $f\colon B^* \to A^*$, $\varphi \circ f(u) \leq \varphi \circ f(v)$.

A monoid (ordered monoid) M satisfies the identity $u = v$ ($u \leq v$) if for every morphism $\varphi\colon B^* \to M$, $\varphi(u) = \varphi(v)$ ($\varphi(u) \leq \varphi(v)$).

An lm-variety $\mathbf{V}$ satisfies a given lm-identity if every stamp in $\mathbf{V}$ satisfies this identity. The class of all stamps satisfying a given set of lm-identities is an lm-variety of stamps. Similarly the class of all (ordered) monoids satisfying a given set of identities is an variety of (ordered) monoids.

By extension, we say that a language L satisfies a monoid identity (lm-identity) if its syntactic monoid (ordered monoid, stamp, ordered stamp) satisfies this identity.

Example 2.8 As an lm-variety of stamps, **MOD** is defined by the single identity $x^{\omega-1}y = 1$.

The variety of finite aperiodic monoids **A** is defined by the identity $x^\omega = x^{\omega+1}$.

The variety of finite ordered monoids $\mathbf{J}^+$ is defined by the identity $x \leq 1$. The variety of finite monoids **J** is defined by the two identities $x^\omega = x^{\omega+1}$ and $(xy)^\omega = (yx)^\omega$.

3. Expressive power of $\mathbf{\Sigma_1}[< + \text{MOD}]$

We first give a simple combinatorial description of the languages definable in $\mathbf{\Sigma_1}[< + \text{MOD}]$.

Let us call *modular simple* a language of the form $(A^d)^*a_1(A^d)^*a_2(A^d)^* \cdots\ a_k(A^d)^*$, where $d > 0$, $k \geq 0$ and $a_1, a_2, \ldots, a_k \in A$.

Proposition 3.1 *A language is definable in* $\mathbf{\Sigma_1}[<+\text{MOD}]$ *if and only if it is a finite union of modular simple languages.*

Proof. The language $(A^d)^*a_1(A^d)^*a_2(A^d)^* \cdots\ a_k(A^d)^*$ can be defined by the Σ_1-formula

$$\exists x_1\ \ldots\ \exists x_k\ (x_1 < \ldots < x_k) \wedge (\mathbf{a}_1 x_1 \wedge \cdots \wedge \mathbf{a}_k x_k) \wedge (\text{MOD}^d_0\, x_1 \wedge \text{MOD}^d_1\, x_2 \wedge \cdots \wedge \text{MOD}^d_{k-1}\, x_k \wedge \text{MOD}^d_{k-1}\, m)$$

This shows that any finite union of modular simple languages is definable in $\mathbf{\Sigma_1}[< + \text{MOD}]$. To prove the result in the opposite direction, consider a Σ_1-formula $\psi = \exists x_1\ \ldots\ \exists x_k\ \varphi(x_1, \ldots, x_k)$. We may assume that φ is in disjunctive normal form. Negations of atomic formulas can be eliminated by replacing $\neg(x = y)$ by $(x < y) \vee (y < x)$, $\neg(x < y)$ by $(x = y) \vee (y < x)$, $\neg(\text{MOD}^d_r\, x)$ by $\vee_{s \neq r} \text{MOD}^d_s\, x$ and $\neg(\mathbf{a}x)$ by $\vee_{b \neq a}(\mathbf{b}x)$. Further, by the Chinese remainder theorem, conjunctions of atomic formulas of the form $\text{MOD}^{d_0}_{r_0}\, m \wedge \bigwedge_{1 \leq i \leq n} \text{MOD}^{d_i}_{r_i}\, x_i$ can be replaced by disjunctions of formulas of the form $\text{MOD}^d_{s_0}\, m \wedge \bigwedge_{1 \leq i \leq n} \text{MOD}^d_{s_i}\, x_i$, where $d = \text{lcm}(d_i)$. Altogether, ψ is equivalent to a disjunction of formulas of the form $\exists x_1\ \ldots\ \exists x_k\ (x_1 < \ldots < x_k) \wedge (\mathbf{a}_1 x_1 \wedge \cdots \wedge \mathbf{a}_k x_k) \wedge (\text{MOD}^d_{r_1}\, x_1 \wedge \cdots \wedge \text{MOD}^d_{r_k}\, x_k \wedge \text{MOD}^d_r\, m)$ defining the language $(A^d)^*A^{s_1}a_1(A^d)^*A^{s_2}a_2(A^d)^* \cdots\ a_k(A^d)^*A^s$ where, for $1 \leq i \leq k$, $s_1 + s_2 + \cdots + s_i \equiv r_i \pmod d$ and $r_k + s \equiv r \pmod d$. Finally, observing that $(A^d)^*A^r = [(A^d)^*(\cup_{a \in A} a)]^r(A^d)^*$, it suffices to use the distributivity of concatenation over union to conclude that the language $L(\psi)$ is a finite union of modular simple languages. □

The concatenation hierarchy of star-free languages mentioned in the introduction is defined by alternating two types of operations: the Boolean operations and the polynomial closure, that we now define. Given a class of languages $\mathcal{L}$, we denote by Pol($\mathcal{L}$) the *polynomial closure* of $\mathcal{L}$, which is the class of languages that are finite unions of languages of the form $L_0a_1L_1a_2 \cdots a_kL_k$, where $L_0, \ldots, L_k \in \mathcal{L}$ and $a_1, \ldots, a_k$ are letters. We also denote by $\mathcal{B}$Pol($\mathcal{L}$) the Boolean closure of Pol($\mathcal{L}$).

It is shown in [13] that $\mathbf{\Sigma_1}[<]$ is equal to Pol($\mathcal{I}$) where $\mathcal{I}$ is the trivial variety of languages. The next proposition shows that $\mathbf{\Sigma_1}[< + \text{MOD}]$ is equal to Pol($\mathcal{M}od$).

Proposition 3.2 *A language belongs to Pol($\mathcal{M}od$) if and only if it is a finite union of modular simple languages.*

Proof. First, Pol($\mathcal{M}od$) clearly contains the modular simple languages. Conversely, any language of Pol($\mathcal{M}od$)(A^*) can be written as a finite union of languages of the form $L = L_0a_1L_1a_2 \cdots a_kL_k$, where $a_1, \ldots, a_k$ are letters and $L_0, \ldots, L_k \in \mathcal{M}od(A^*)$. Thus each L_i is a finite union of languages of the form $(A^{n_i})^*A^k$, with $0 \leq k \leq n_i$. Let d be the least common multiple of the n_i. Setting $r_i = d/n_i$, we observe that $(A^{n_i})^* = \cup_{0 \leq k < r_i}(A^d)^*A^{kn_i}$. Applying the distributivity of concatenation over union, we may assume that all L_i are of the form $(A^d)^*A^k$. But $(A^d)^*A^k$ can be written as $\cup_{a_1a_2 \cdots a_k \in A^k}(A^d)^*a_1(A^d)^*a_2 \cdots a_k(A^d)^*$. It follows that any language of Pol($\mathcal{M}od$)(A^*) is a finite union of modular simple languages. □

Our decidability result for $\mathbf{\Sigma_1}[< + \text{MOD}]$ relies on an algebraic characterization of the polynomial closure [16, 17]. However, the formulation of this general result requires us to introduce Mal'cev products of varieties and we prefer here a simpler formulation.

Proposition 3.3 *A language belongs to Pol($\mathcal{M}od$) if and only if its ordered syntactic stamp φ satisfies the following property: there exists a positive integer n such that the ordered monoid $\varphi((A^n)^*)$ satisfies the identity $x \leq 1$.*

Unfortunately, Proposition 3.3 does not provide a decidability criterion for Pol($\mathcal{M}od$). The next result fixes this problem.

Theorem 3.4 *A language belongs to Pol($\mathcal{M}od$) if and only if the stable ordered monoid of its ordered syntactic stamp satisfies the identity $x \leq 1$.*

Proof. By Proposition 3.3, it suffices to show that if $\varphi((A^n)^*)$ satisfies the identity $x \leq 1$ for some $n > 0$, then $\varphi((A^s)^*)$ satisfies the same identity. But since $\varphi(A^s) =$

$\varphi(A^{ns})$, $\varphi((A^s)^*) = \varphi((A^{ns})^*)$. It follows that $\varphi((A^s)^*)$ is a submonoid of $\varphi((A^n)^*)$ and thus satisfies the identity $x \leq 1$. $\square$

Theorem 3.4 gives a decidable condition for testing membership in Pol($\mathcal{M}od$). But since we know that Pol($\mathcal{M}od$) is a positive lm-variety of languages, it is interesting to find the identities defining the corresponding variety of ordered stamps.

Theorem 3.5 *A language belongs to Pol($\mathcal{M}od$) if and only if its ordered syntactic stamp satisfies the lm-identities $x^{\omega-1}y \leq 1$ and $yx^{\omega-1} \leq 1$.*

Proof. Let L be a regular language, $\varphi\colon A^* \to M$ its ordered syntactic stamp, S its stable monoid and s its stability index.

First assume that L belongs to Pol($\mathcal{M}od$). Let x and y be two words in A^* of equal length and let $u = x^{(s-1)\omega}x^{\omega-1}y$. The length of u is a multiple of s and thus $\varphi(u)$ belongs to S. By Theorem 3.4, S satisfies the identity $x \leq 1$ and hence $\varphi(u) \leq 1$. But $\varphi(u) = \varphi(x^{\omega-1}y)$ and thus $\varphi(x^{\omega-1}y) \leq 1$. This proves that φ satisfies the lm-identities $x^{\omega-1}y \leq 1$. A symmetrical argument works for the second identity.

Conversely, assume that φ satisfies the lm-identities $x^{\omega-1}y \leq 1$ and $yx^{\omega-1} \leq 1$. We claim that $m \leq 1$ for all $m \in S$. The relation is trivial if $m = 1$. If $m \neq 1$, then $m \in \varphi(A^s) = T$. Since $T^2 = T$, it follows from [14, Chap. 1, Proposition 1.12] that $m = uev$ for some $u, e, v \in T$ with e idempotent. Thus there exist $x, y, z \in A^s$ such that $\varphi(y) = u$, $\varphi(x) = e$ and $\varphi(z) = v$. Since $|x| = |y| = |z|$, one has $\varphi(yx^{\omega-1}) \leq 1$ and $\varphi(x^{\omega-1}z) \leq 1$. It follows that $ue \leq 1$ and $ev \leq 1$, whence $m = uev = ueev \leq 1$. This proves the claim and shows, by Theorem 3.4, that L belongs to Pol($\mathcal{M}od$). $\square$

The results of this section should be compared with the characterization of the class $\mathbf{\Sigma_1}[< + \text{REG}]$ which can be derived from the two papers [8, 21].

4. Expressive power of $\mathcal{B}\mathbf{\Sigma_1}[< + \text{MOD}]$

In this section we give several characterizations of the class $\mathcal{B}\mathbf{\Sigma_1}[< + \text{MOD}]$. Let us start with an immediate consequence of Proposition 3.1:

Proposition 4.1 *A language is definable in $\mathcal{B}\mathbf{\Sigma_1}[<+\text{MOD}]$ if and only if it is a Boolean combination of modular simple languages.*

Our second characterization is based on properties of the wreath product. The non-specialist reader can skip the technical definitions given below, admit Theorem 4.2 and jump directly to Theorem 4.3.

The wreath product $N \circ K$ of two monoids N and K is defined on the set $N^K \times K$ by the following product:

$$(f_1, k_1)(f_2, k_2) = (f, k_1k_2), \text{ with } f(k) = f_1(k)f_2(kk_1)$$

This definition can be extended to varieties of stamps as follows. Let $\mathbf{V}, \mathbf{W}$ be two lm-varieties of stamps. A $(\mathbf{V}, \mathbf{W})$-product stamp is a stamp $\varphi : A^* \to M$ such that:

(1) M is a submonoid of a wreath product $N \circ K$, where N and K are finite monoids.

(2) Let $\pi : N \circ K \to K$ be the canonical projection morphism. Then the stamp $\pi \circ \varphi : A^* \to \pi(M)$ is in $\mathbf{W}$.

(3) For a in A, we can write $\varphi(a) = (f_a, \pi \circ \varphi(a))$ where f_a is in N^K. We now treat $K \times A$ as a finite alphabet and we define a stamp $\Phi : (K \times A)^* \to \text{Im}(\Phi) \subseteq N$ by $\Phi(k, a) = f_a(k)$. We require Φ to be in $\mathbf{V}$.

We define $\mathbf{V} * \mathbf{W}$ to be the class of all stamps that divide a $(\mathbf{V}, \mathbf{W})$-product stamp. The class $\mathbf{V} * \mathbf{W}$ is called the *wreath product* of the lm-varieties of stamps $\mathbf{V}$ and $\mathbf{W}$. It can be shown [3] that $\mathbf{V} * \mathbf{W}$ is an lm-variety of stamps containing $\mathbf{W}$. The wreath product is an associative operation on lm-varieties of stamps which extends the classical wreath product on Eilenberg's varieties.

The wreath product principle [6, 3] gives a description of languages recognized by a stamp of $\mathbf{V} * \mathbf{W}$. It is based on similar results for varieties of monoids [24, 18]. We only give here a simplified version for the case $\mathbf{W} = \mathbf{MOD}$.
For each $n > 0$, let $B_n = \mathbb{Z}/n\mathbb{Z} \times A$ and $\sigma_n : A^* \to B_n^*$ be the sequential function defined by setting:

$$\sigma_n(a_1 \cdots a_k) = (0, a_1)(1, a_2) \cdots (k-1, a_k).$$

Theorem 4.2 *Let $\mathbf{V}$ be an lm-variety of stamps and let $\mathcal{U}$ be the lm-variety of languages associated with $\mathbf{V} * \mathbf{MOD}$. Then for every alphabet A, $\mathcal{U}(A^*)$ is the smallest positive Boolean algebra containing $\mathcal{M}od(A^*)$ and the languages of the form $\sigma_n^{-1}(V)$, where $n > 0$ and V is in $\mathcal{V}(B_n^*)$.*

Proof. The general Wreath Product Principle on stamps (WPP for short) [3] makes use of slightly more involved sequential functions that we shall introduce now. Given a stamp $\varphi : A^* \to M$ and an element m in M, we define the sequential function $\rho_m : A^* \to (M \times A)^*$ by setting:

$$\rho_m(a_1 \cdots a_n) = (m, a_1)(m\varphi(a_1), a_2) \cdots (m\varphi(a_1 \cdots a_{n-1}), a_n)$$

A sequential function ρ is said to be *associated with* φ if $\rho = \rho_m$ for some m in M. The WPP states that $\mathcal{U}(A^*)$ is the smallest positive Boolean algebra containing $\mathcal{M}od(A^*)$ and the languages of the form $\rho^{-1}(V)$, where ρ is a sequential function associated with a stamp $\varphi : A^* \to M$ in $\mathbf{MOD}$ and V is in $\mathcal{V}\big((M \times A)^*\big)$.

Notice first that, if $\varphi : A^* \to M$ is in **MOD** then M is a finite cyclic group, and one can thus assume that $M = \mathbb{Z}/n\mathbb{Z}$ for some positive integer n. We denote this group additively. Further, since φ is surjective, there exists a generator k of $\mathbb{Z}/n\mathbb{Z}$ such that $\varphi(A) = \{k\}$. Thus φ is isomorphic to the stamp $\pi_n : A^* \to \mathbb{Z}/n\mathbb{Z}$, defined by $\pi_n(A) = \{1\}$. Therefore $\mathcal{U}(A^*)$ is the smallest positive Boolean algebra containing $\mathcal{M}od(A^*)$ and the languages of the form $\rho^{-1}(V)$, where ρ is a sequential function associated with some stamp π_n and V is in $\mathcal{V}(B_n^*)$.

Now, let V be a language in $\mathcal{V}(B_n^*)$ and let $\rho_k : A^* \to B_n^*$ be the sequential function associated with π_n and an element k in $\mathbb{Z}/n\mathbb{Z}$. Define the lm-morphism $f_k \colon B_n^* \to B_n^*$ by $f_k(x, a) = (x + k, a)$, and let $V' = f_k^{-1}(V)$. Then V' is in $\mathcal{V}(B_n^*)$ and $\rho_k^{-1}(V) = \sigma_n^{-1}(V')$. Therefore, it is sufficient to consider sequential functions of the form σ_n, which concludes the proof. $\square$

We now arrive at our second characterization of $\mathcal{B}\mathbf{\Sigma_1}[< + \textsc{mod}]$.

Theorem 4.3 *A language is a Boolean combination of modular simple languages if and only if its syntactic stamp belongs to the lm-variety* $\mathbf{J} * \mathbf{MOD}$.

Proof. Let $\mathcal{U}$ be the lm-variety of languages corresponding to $\mathbf{J} * \mathbf{MOD}$. We first show that each language of $\mathcal{U}$ is a Boolean combination of modular simple languages. By Proposition 3.2, it suffices to show that $\mathcal{U}$ is contained in $\mathcal{B}\mathrm{Pol}(\mathcal{M}od)$.

Let A be an alphabet. According to Theorem 4.2, $\mathcal{U}(A^*)$ is the smallest positive Boolean algebra containing $\mathcal{M}od(A^*)$ and the languages of the form $\sigma_n^{-1}(V)$, where $n > 0$ and V belongs to $\mathcal{J}(B_n^*)$. Since $\mathcal{M}od$ is contained in $\mathrm{Pol}(\mathcal{M}od)$, it remains to prove that all languages of the form $\sigma_n^{-1}(V)$ are in $\mathcal{B}\mathrm{Pol}(\mathcal{M}od)$. Further, since σ_n^{-1} commutes with Boolean operations, we may assume by Simon's theorem [22] that V is equal to $B_n^* b_1 B_n^* \cdots b_p B_n^*$ for some $b_1, \ldots, b_p \in B_n$. Setting $b_i = (r_i, a_i)$, we observe that

$$\sigma_n^{-1}(V) = (A^n)^* A^{r_1} a_1 (A^n)^* A^{s_2} a_2 \cdots (A^n)^* A^{s_p} a_p A^*,$$

with $s_i = r_i - (r_{i-1} + 1) \bmod n$, for $i = 2 \cdots p$. Since A^* and all languages of the form $(A^n)^* A^j$ are in $\mathcal{M}od(A^*)$, $\sigma^{-1}(V)$ belongs to $\mathrm{Pol}(\mathcal{M}od(A^*))$.

We now prove that any Boolean combination of modular simple languages is in $\mathcal{U}$. A simple computation shows that if

$$L = (A^d)^* a_1 (A^d)^* a_2 (A^d)^* \cdots\ a_k (A^d)^*$$

is a modular simple language of A^*, then

$$L = \sigma_d^{-1}(B_d^* b_1 B_d^* \cdots b_k B_d^*) \cap (A^d)^* A^k$$

with $b_i = (i-1, a_i)$ for $1 \leq i \leq k$. Since $B_d^* b_1 B_d^* \cdots b_k B_d^*$ is in $\mathcal{J}(B_d^*)$, L belongs to $\mathcal{U}(A^*)$. Finally, since $\mathcal{U}(A^*)$ is a Boolean algebra, any Boolean combination of modular simple languages of A^* is in $\mathcal{U}(A^*)$. $\square$

It follows from Proposition 4.1 and Theorem 4.3 that deciding whether a given regular language is definable in $\mathcal{B}\mathbf{\Sigma_1}[< + \textsc{mod}]$ amounts to showing that the lm-variety $\mathbf{J} * \mathbf{MOD}$ is decidable. The proof requires us to introduce derived categories [28]. In this paper, categories are viewed as generalizations of monoids since a one-object category is in fact a monoid.

Let C, D be two categories. A *division* of categories $\tau : C \to D$ is given by a mapping $\tau : Obj(C) \to Obj(D)$ and for each pair (u, v) of objects of C, by a relation $\tau : C(u, v) \to D(\tau(u), \tau(v))$ such that

(1) $\tau(x)\tau(y) \subseteq \tau(xy)$ for any consecutive arrows x, y,
(2) $\tau(x) \neq \emptyset$ for any arrow x,
(3) $1_{\tau(u)} \in \tau(1_u)$,
(4) $\tau(x) \cap \tau(y) \neq \emptyset$ implies $x = y$ for any coterminal arrows x, y of C.

If $\mathbf{V}$ is variety of monoids, we denote by $g\mathbf{V}$ the class of all categories that divide a monoid in $\mathbf{V}$ (regarded as a one-object category). By transitivity of division of categories, $g\mathbf{V}$ is always closed under division.

Let $\varphi : A^* \to M$ be a stamp. For each integer n, let $\pi_n : A^* \to \mathbb{Z}/n\mathbb{Z}$ be the stamp defined by $\pi_n(u) = |u| \bmod n$ and let φ_n be the relational morphism $\varphi_n = \pi_n \circ \varphi^{-1}$.

$$\begin{tikzcd} & A^* \arrow[dl, "\varphi"'] \arrow[dr, "\pi_n"] & \\ M \arrow[rr, "\varphi_n"] & & \mathbb{Z}/n\mathbb{Z} \end{tikzcd}$$

Let $C_n(\varphi)$ be the category whose objects are elements of $\mathbb{Z}/n\mathbb{Z}$ and whose arrows from object i to object j are the triples (i, m, j) where $j - i \in \varphi_n(m)$. Its composition rule is given by $(i, m_1, j)(j, m_2, k) = (i, m_1 m_2, k)$.

The next result is a special instance of the derived category theorem due to Tilson [28], but two modifications occur. First, Tilson's original definition of the derived category was different from ours, but this more complex definition is not required for relational morphisms onto a group. Second, Tilson's proof needs to be adapted to the context of stamps. Altogether, we obtain the following result:

Theorem 4.4 *A stamp φ is in* $\mathbf{J} * \mathbf{MOD}$ *if and only if there exists a positive integer n such that $C_n(\varphi)$ is in $g\mathbf{J}$.*

We shall now improve Theorem 4.4 by giving an explicit bound on the integer n. First, it was shown by Knast that a category belongs to $g\mathbf{J}$ if and only if, for each of its subgraphs of the form given in Figure 2, one has

$$(m_1 m_2)^\omega (m_3 m_4)^\omega = (m_1 m_2)^\omega m_1 m_4 (m_3 m_4)^\omega \qquad (1)$$

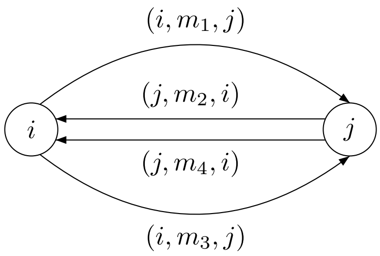


Figure 2. A Knast subgraph.

We now state our new characterization.

Theorem 4.5 *Let φ be a stamp of stability index s. Then φ belongs to* $\mathbf{J} * \mathbf{MOD}$ *if and only if $C_s(\varphi)$ is in $g\mathbf{J}$.*

Proof. First, if $C_s(\varphi)$ is in $g\mathbf{J}$, then φ belongs to $\mathbf{J} * \mathbf{MOD}$ by Theorem 4.4.

Now assume that $\varphi\colon A^* \to M$ belongs to $\mathbf{J} * \mathbf{MOD}$. Then, by Theorem 4.4, there exists a positive integer n such that $C_n(\varphi)$ is in $g\mathbf{J}$. We prove that $C_s(\varphi)$ is in $g\mathbf{J}$ by showing that it satisfies Knast's equation. Consider a Knast subgraph of $C_s(\varphi)$, with the notation in Figure 2. Set $k = j - i$. There exist words u_1, u_2, u_3, u_4 in A^* such that $\varphi(u_i) = m_i$ for $1 \leq i \leq 4$ and

$$|u_1| \equiv |u_3| \equiv -|u_2| \equiv -|u_4| \equiv k \bmod s.$$

Since M is a finite monoid, there exists an integer ω such that, for all $x \in M$, x^ω is idempotent. Further we can assume that ω is greater than s. Now setting

$$\begin{cases} v_1 = (u_1u_2)^\omega u_1, & v_2 = u_2(u_1u_2)^{\omega-1} \\ v_3 = (u_3u_4)^\omega u_3, & v_4 = u_4(u_3u_4)^{\omega-1} \end{cases}$$

we still have $|v_1| \equiv |v_3| \equiv -|v_2| \equiv -|v_4| \equiv k \bmod s$. Further $(\varphi(v_1), \varphi(v_2))$ and $(\varphi(v_3), \varphi(v_4))$ are pairs of mutually inverse elements of M. If $k \neq 0$, then for each i, $|v_i| \geq s$ and one can find an integer p_i such that

$$\begin{cases} |v_i| = p_is + k\ , \ \ p_i > 0, & \text{for } i = 1, 3 \\ |v_i| = p_is - k\ , \ \ p_i > 1, & \text{for } i = 2, 4 \end{cases}$$

By definition of s, we have $\varphi(A^s) = \varphi(A^{2s})$ and hence

$$\begin{cases} \varphi(A^{p_is+k}) = \varphi(A^{np_is+k}), & \text{for } i = 1, 3 \\ \varphi(A^{p_is-k}) = \varphi(A^{np_is-k}), & \text{for } i = 2, 4 \end{cases}$$

Thus, there exist words x_1, x_2, x_3, x_4 in A^* such that $\varphi(v_i) = \varphi(x_i)$ for $1 \leq i \leq 4$ and

$$\begin{cases} |x_i| = np_is + k, & \text{for } i = 1, 3 \\ |x_i| = np_is - k, & \text{for } i = 2, 4 \end{cases}$$

Therefore, $|x_1| \equiv |x_3| \equiv -|x_2| \equiv -|x_4| \equiv k \bmod n$, and $C_n(\varphi)$ contains the subgraph pictured in Figure 3.

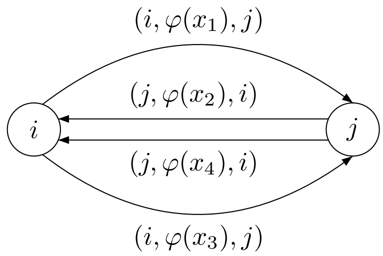


Figure 3. A subgraph of $C_n(\varphi)$.

Since $C_n(\varphi)$ is in $g\mathbf{J}$, it satisfies Knast's equation, that is,

$$\varphi(x_1x_2)^\omega\varphi(x_3x_4)^\omega = \varphi(x_1x_2)^\omega\varphi(x_1x_4)\varphi(x_3x_4)^\omega,$$

which finally yields Equation (1). Therefore, $C_s(\varphi)$ is in $g\mathbf{J}$.

We now treat the case where $k = 0$. If $u_1 = u_2 = u_3 = u_4 = 1$, Equation (1) holds trivially. Else, if $u_1 = u_2 = 1$ but $u_3u_4 \neq 1$, we set $x_1 = x_2 = 1$ and since $|v_3|, |v_4| \geq s$, we can take x_3, x_4 as above. Then, it is still true that $|x_1| \equiv |x_3| \equiv -|x_2| \equiv -|x_4| \equiv 0 \bmod n$ and that $C_n(\varphi)$ contains the subgraph pictured in Figure 3, which gives the result. The argument is symmetrical if $u_3 = u_4 = 1$. In all remaining cases, the words v_i have length greater or equal to s and the proof of the case $k \neq 0$ carries over. □

Corollary 4.6 *Given a regular language L, one can effectively decide whether L is definable in $\mathcal{B}\mathbf{\Sigma_1}[< + \text{MOD}]$.*

Proof. It suffices to compute the syntactic stamp of L and its stability index s and check whether the derived category $C_s(\varphi)$ satisfies Knast's identity (1). □

5. Summary

We proved the decidability of the two classes $\mathbf{\Sigma_1}[< + \text{MOD}]$ and $\mathcal{B}\mathbf{\Sigma_1}[< + \text{MOD}]$. In algebraic terms, our results can be summarized as follows:

	$<$	$< + \text{MOD}$	$< + \text{REG}$
$\mathbf{\Sigma_1}$	$\mathbf{J}^+$	$\mathbf{J}^+ * \mathbf{MOD}$	$\mathbf{J}^+ * \mathbf{LI} * \mathbf{MOD}$
$\mathcal{B}\mathbf{\Sigma_1}$	$\mathbf{J}$	$\mathbf{J} * \mathbf{MOD}$	$\mathbf{J} * \mathbf{LI} * \mathbf{MOD}$
$\vdots$			
FO	$\mathbf{A}$	$\mathbf{A} * \mathbf{MOD}$	$\mathbf{A} * \mathbf{MOD}$

However, there are subtle differences between these two new results, as well as important features that distinguish them from the older results listed in the fourth column of the table. Indeed, given a stamp φ, one can decide whether φ belongs to the varieties of the fourth column by verifying that their stable (ordered) monoid satisfies certain conditions. This is due to the properties that the varieties $\mathbf{J}^+ * \mathbf{LI}$, $\mathbf{J} * \mathbf{LI}$ and $\mathbf{A}$ satisfy the condition $\mathbf{V} * \mathbf{LI} = \mathbf{V}$. It was observed both in [6] and in [12] that for varieties satisfying this condition, the decidability of $\mathbf{V}$ and $\mathbf{V} * \mathbf{MOD}$ are equivalent. The variety of ordered monoids $\mathbf{J}^+$ does not satisfy this condition, but it is a *local variety* in the sense of Tilson [28]: this still suffices to get the decidability of $\mathbf{J}^+ * \mathbf{MOD}$. The hardest case is $\mathbf{J} * \mathbf{MOD}$: the variety $\mathbf{J}$ is known to be nonlocal and Knast identities are required to get the decidability.

It would be interesting to obtain a purely model theoretic proof of our results.

References

[1] Mustapha Arfi. Opérations polynomiales et hiérarchies de concaténation. *Theoret. Comput. Sci.*, 91(1):71–84, 1991.

[2] David A. Mix Barrington, Kevin J. Compton, Howard Straubing, and Denis Thérien. Regular languages in NC^1. *J. Comput. Syst. Sci.*, 44:478–499, 1992.

[3] Laura Chaubard, Jean-Éric Pin, and Howard Straubing. Actions, wreath products of $\mathcal{C}$-varieties and concatenation product. *Theoret. Comput. Sci.*, 2006. to appear.

[4] Samuel Eilenberg. *Automata, languages, and machines. Vol. B*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. With two chapters (“Depth decomposition theorem” and “Complexity of semigroups and morphisms”) by Bret Tilson, Pure and Applied Mathematics, Vol. 59.

[5] Zoltán Ésik. Extended temporal logic on finite words and wreath products of monoids with distinguished generators. In Masami et al. Ito, editor, *Developments in language theory. 6th international conference, DLT 2002, Kyoto, Japan, September 18-21*, number 2450 in Lect. Notes Comp. Sci., pages 43–58, Berlin, 2002. Springer.

[6] Zoltán Ésik and Masami Ito. Temporal logic with cyclic counting and the degree of aperiodicity of finite automata. *Acta Cybernetica*, 16:1–28, 2003.

[7] Zoltán Ésik and Kim G. Larsen. Regular languages definable by lindström quantifiers. *Theor. Inform. Appl.*, 37:179–241, 2003.

[8] Christian Glaßer. Polylog-time reductions decrease dot-depth. In *STACS 2005*, volume 3404 of *Lect. Notes Comp. Sci.*, pages 170–181. Springer, Berlin, 2005.

[9] Michal Kunc. Equational description of pseudovarieties of homomorphisms. *Theoretical Informatics and Applications*, 37:243–254, 2003.

[10] Alexis Maciel, Pierre Péladeau, and Denis Thérien. Programs over semigroups of dot-depth one. *Theoret. Comput. Sci.*, 245(1):135–148, 2000.

[11] Robert McNaughton and Seymour Papert. *Counter-free automata*. The M.I.T. Press, Cambridge, Mass.-London, 1971. With an appendix by William Henneman, M.I.T. Research Monograph, No. 65.

[12] Pierre Péladeau, Howard Straubing, and Denis Thérien. Finite semigroup varieties defined by programs. *Theoret. Comput. Sci.*, 180(1-2):325–339, 1997.

[13] Dominique Perrin and Jean-Éric Pin. First order logic and star-free sets. *J. Comput. System Sci.*, 32:393–406, 1986.

[14] Jean-Éric Pin. *Varieties of formal languages*. North Oxford, London and Plenum, New-York, 1986. (Traduction de Variétés de langages formels).

[15] Jean-Éric Pin. A variety theorem without complementation. *Russian Mathematics (Izvestija vuzov.Matematika)*, 39:80–90, 1995.

[16] Jean-Éric Pin and Howard Straubing. Some results on $\mathcal{C}$-varieties. *Theoret. Informatics Appl.*, 39:239–262, 2005.

[17] Jean-Éric Pin and Pascal Weil. Polynomial closure and unambiguous product. *Theory Comput. Systems*, 30:1–39, 1997.

[18] Jean-Éric Pin and Pascal Weil. The wreath product principle for ordered semigroups. *Communications in Algebra*, 30:5677–5713, 2002.

[19] Jan Reiterman. The Birkhoff theorem for finite algebras. *Algebra Universalis*, 14(1):1–10, 1982.

[20] Marcel-Paul Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8:190–194, 1965.

[21] Victor L. Selivanov. Some hierarchies and reducibilities on regular languages. Technical Report 349, University of Würzburg, Germany, 2004.

[22] Imre Simon. *Hierarchies of Events with Dot-Depth One*. PhD thesis, University of Waterloo, Waterloo, Ontario, Canada, 1972.

[23] Imre Simon. Piecewise testable events. In H. Brackage, editor, *Proc. 2nd GI Conf.*, volume 33 of *Lecture Notes in Comp. Sci.*, pages 214–222. Springer Verlag, Berlin, Heidelberg, New York, 1975.

[24] Howard Straubing. Families of recognizable sets corresponding to certain varieties of finite monoids. *J. Pure Appl. Algebra*, 15(3):305–318, 1979.

[25] Howard Straubing. *Finite automata, formal logic, and circuit complexity*. Birkhäuser Boston Inc., Boston, MA, 1994.

[26] Howard Straubing. On logical descriptions of regular languages. In *LATIN 2002*, number 2286 in Lect. Notes Comp. Sci., pages 528–538, Berlin, 2002. Springer.

[27] Wolfgang Thomas. Classifying regular events in symbolic logic. *J. Comput. System Sci.*, 25(3):360–376, 1982.

[28] Bret Tilson. Categories as algebra: an essential ingredient in the theory of monoids. *J. Pure Appl. Algebra*, 48(1-2):83–198, 1987.