

Separation and the Successor Relation*

Thomas Place and Marc Zeitoun

LaBRI, Bordeaux University, France, firstname.lastname@labri.fr.

Abstract

We investigate two problems for a class \mathcal{C} of regular word languages. The \mathcal{C} -membership problem asks for an algorithm to decide whether an input language belongs to \mathcal{C} . The \mathcal{C} -separation problem asks for an algorithm that, given as input two regular languages, decides whether there exists a third language in \mathcal{C} containing the first language, while being disjoint from the second. These problems are considered as means to obtain a deep understanding of the class \mathcal{C} .

It is usual for such classes to be defined by logical formalisms. Logics are often built on top of each other, by adding new predicates. A natural construction is to enrich a logic with the successor relation. In this paper, we obtain new and simple proofs of two transfer results: we show that for suitable logically defined classes, the membership, resp. the separation problem for a class enriched with the successor relation reduces to the same problem for the original class.

Our reductions work both for languages of finite words and infinite words. The proofs are mostly self-contained, and only require a basic background on regular languages. This paper therefore gives simple proofs of results that were considered as difficult, such as the decidability of the membership problem for the levels 1, $3/2$, 2 and $5/2$ of the dot-depth hierarchy.

1998 ACM Subject Classification F.4.3 Formal Languages

Keywords and phrases separation problem, regular word languages, logics, decidable characterizations, semidirect product

Digital Object Identifier 10.4230/LIPIcs.STACS.2015.662

1 Introduction

A central problem in the theory of formal languages is to characterize and understand the expressive power of high level specification formalisms. Monadic second order logic (MSO) is such a formalism, which is both expressive and robust. For several classes of structures, such as words or trees, it has the same expressive power as finite automata and defines the class of regular languages. In this paper, we investigate fragments of MSO over words. In this context, understanding the expressive power of a fragment is associated to two decision problems: the *membership problem* and the *separation problem*.

For a fixed logical fragment \mathcal{F} , the \mathcal{F} -*membership problem* asks for a decision procedure that tests whether some input regular language can be expressed by a formula from \mathcal{F} . To obtain such an algorithm, one has to consider and understand *all* properties that can be expressed within \mathcal{F} , which requires a deep understanding of the fragment \mathcal{F} . On the other hand, the \mathcal{F} -*separation problem* is more general. It asks for a decision procedure that tests whether given *two* input regular languages, there exists a third one in \mathcal{F} containing the first language while being disjoint from the second one.

Since regular languages are closed under complement, membership reduces to separation: a language is in \mathcal{F} if and only if it can be separated from its complement. Usually, the

* Supported by ANR 2010 BLAN 0202 01 FREC

separation problem is more difficult than the membership problem but also more rewarding with respect to the knowledge gained on the investigated fragment \mathcal{F} .

These two problems have been considered and solved for many natural fragments of monadic second order logic. Among these, the most prominent one is first-order logic, $\text{FO}(<)$, equipped with a predicate $<$ for the linear ordering. The solution to the membership problem, known as the McNaughton-Papert-Schützenberger Theorem [20, 10], has been revisited until recently [5]. The theorem states that a regular language is definable in $\text{FO}(<)$ if and only if its *syntactic semigroup* is aperiodic. The syntactic semigroup is a finite algebraic object that can be computed from any regular language. Since aperiodicity can be defined as an equation that needs to be satisfied by all of its elements, this yields decidability of $\text{FO}(<)$ -definability. This result now serves as a template, which is commonly followed in this line of research.

The separation problem has also been successfully solved for first-order logic [7]. Actually, the problem was first addressed in a purely algebraic framework, and was later identified as equivalent to our separation problem [2]. As for membership, this problem is still revisited today and a new self-contained and combinatorial proof was obtained in [18].

Motivation. We are interested in natural fragments of $\text{FO}(<)$ obtained by restricting either the number of variables or the number of quantifier alternations allowed in formulas. Such restrictions in general give rise to several variants of the same fragment. Indeed, in most cases, the drop in expressive power forbids the use of natural relations that could be defined from the linear order in $\text{FO}(<)$. The main example considered in this paper is $+1$, the *successor relation*, together with predicates *min* and *max* for the first and last positions in a word. This means that one can define two distinct variants of the same fragment depending on whether we decide to explicitly add these predicates in the signature or not. An example is the fragment Σ_n , which consists of first-order formulas whose prenex normal form has at most $(n - 1)$ quantifier alternations and starts with an existential block. Since defining $+1$ requires an additional quantifier alternation, $\Sigma_n(<, +1, \text{min}, \text{max})$ has indeed stronger expressiveness than $\Sigma_n(<)$. The motivation of this paper is to obtain decidability results for such enriched fragments.

State of the Art. Even when the weak fragment is known to have decidable membership, proving that the enriched one has the same property can be nontrivial. Examples include the membership proofs of $\mathcal{B}\Sigma_1(<, +1, \text{min}, \text{max})$ (Boolean combinations of $\Sigma_1(<, +1, \text{min}, \text{max})$ formulas) and $\Sigma_2(<, +1)$, which require difficult and intricate combinatorial arguments [8, 6, 9] or a wealth of algebraic machinery [12, 13]. Another issue is that most proofs directly deal with the enriched fragment. Given the jungle of such logical fragments, it is desirable to avoid such an approach, treating each variant of the same fragment independently. Instead, a satisfying approach is to first obtain a solution of the membership and separation problems for the less expressive variant and then to lift it to other variants via a generic transfer result.

This approach has first been investigated by Straubing for the membership problem [22] in an algebraic framework, and later adapted to be able to treat classes not closed under complement [13]. Transferring the logical problem to this algebraic framework requires preliminary steps, still *specific* to the investigated class, to prove that:

1. A language is definable in the fragment if and only if its syntactic semigroup belongs to a specific algebraic variety \mathcal{V} (e.g., the variety of aperiodic monoids for $\text{FO}(<)$), and
2. Membership to \mathcal{V} is decidable.

Next, though this is not immediate, for most fragments of $\text{FO}(<)$, it has been proved that

3. When the weaker variant corresponds to a variety \mathcal{V} , the variant with successor corresponds to the variety $\mathcal{V} * \mathcal{D}$, built generically from \mathcal{V} .

Hence, Straubing's approach was to prove that

4. the operator $\mathcal{V} \mapsto \mathcal{V} * \mathcal{D}$ preserves decidability.

Unfortunately, this is not true in general [3]. Actually, while decidability is preserved for all known logical fragments, there is no generic result that captures them all. In particular, for the less expressive fragments, one has to use completely *ad hoc* proofs. In the separation setting, things behave well: it has been shown that decidability of separation is preserved by the operation $\mathcal{V} \mapsto \mathcal{V} * \mathcal{D}$ [21]. While interesting when already starting from algebra, this approach has several downsides:

- Dealing with algebra hides the logical intuitions, while our primary goal is to understand the expressiveness of logics.
- Going from logic to algebra requires to be acquainted with new notions and vocabulary, as well as involved theoretical tools. Proofs are also often nontrivial and require a deep understanding of complex objects, which may be scattered in the bibliography.
- Despite step 4, which is generic to some extent, arguments specific to the investigated class are pushed to steps 1–3, and they are often nontrivial.

Contributions. We give a new proof that decidability of separation can be transferred from a weak to an enriched fragment. We present the result in two different forms.

The first one is non-algebraic: we work directly with the logical fragments, without using varieties. The transfer result is generic and its proof mostly is: the only specific argument is an Ehrenfeucht-Fraïssé game that can be adapted to all natural fragments with minimal difficulty (we prove it in the long version of this paper for all considered fragments, see [19]). The benefits of this new proof are that:

1. It is self-contained and much simpler than previous ones. It only relies on two basic well-known notions: recognizability by semigroups and Ehrenfeucht-Fraïssé games.
2. It works with classes that are not closed under complement, contrary to [21]. This allows us to capture the Σ and Π levels in the quantifier alternation hierarchy of first-order logic.
3. Under an additional hypothesis on the logical fragment, which is met for most fragments we investigate and easy to check, the decidability result of the separation problem also extends to the membership problem.
4. The proof adapts smoothly to infinite words using the notion of ω -semigroups, as shown in the long version of this paper [19].

The second form of our result is algebraic and generic. We prove that $\mathcal{V} \mapsto \mathcal{V} * \mathcal{D}$ preserves the decidability of separation for varieties, hence giving an elementary proof of a result of [21]. Even in this algebraic form, we completely bypass involved constructions or notions, such as pointlike sets for categories developed in [21], thus making the proof accessible.

As corollaries, since $\mathcal{B}\Sigma_1(<)$ and $\Sigma_2(<)$ both enjoy decidable separation [4, 16, 17], we obtain that this is also the case for the fragments $\mathcal{B}\Sigma_1(<, +1, \text{min}, \text{max})$ and $\Sigma_2(<, +1)$, known as levels 1 and 3/2 of the dot-depth hierarchy. These new results strengthen the previous ones [8, 6] that showed decidability of membership and were considered as difficult. We actually obtain that separation for $\Sigma_n(<, +1, \text{min}, \text{max})$ reduces to separation for $\Sigma_n(<)$. Since we also transfer decidability of the membership problem, and since the fragments $\mathcal{B}\Sigma_2(<)$ of Boolean combinations of $\Sigma_2(<)$ formulas and $\Sigma_3(<)$ have decidable membership [17] we

deduce that the same holds for $\mathcal{B}\Sigma_2(<, +1)$ and $\Sigma_3(<, +1)$, known as levels 2 and 5/2 of the dot-depth hierarchy.

Organization of the Paper. In Section 2, we set up the notation and we present the separation problem and the logics we deal with. Section 3 is devoted to our main tool: languages of well-formed words. In Section 4, we use it to prove our transfer result for all fragments from the logical perspective, and in Section 5, we show that decidability of the separation problem for the variety \mathbb{V} entails the same for $\mathbb{V} * \mathbb{D}$.

2 Preliminaries

In this section, we provide preliminary definitions on regular languages defined by logical fragments and on separation. We also present our main contribution.

Words, Languages. We fix a finite alphabet A . Let A^+ be the set of all nonempty finite words and let A^* be the set of all finite words over A . If u, v are words, we denote by $u \cdot v$ or by uv the word obtained by concatenating u and v . For convenience, we only consider, without loss of generality, languages that do not contain the empty word. That is, a language is a subset of A^+ . We work with regular languages, that is, languages definable by finite automata.

Separation. Given three languages K, L, L' , we say that K *separates* L from L' if

$$L \subseteq K \text{ and } K \cap L' = \emptyset.$$

If \mathcal{C} is a class of languages, we say that L is \mathcal{C} -*separable* from L' if there exists $K \in \mathcal{C}$ that separates L from L' . Note that if \mathcal{C} is closed under complement, L is \mathcal{C} -separable from L' if and only if L' is \mathcal{C} -separable from L . However, this is not true for a class \mathcal{C} not closed under complement, such as the classes $\Sigma_n(<)$ of the quantifier alternation hierarchy, which we shall consider.

Given a class \mathcal{C} , the \mathcal{C} -*separation problem* asks for an algorithm which, given as input two regular languages L, L' , decides whether L is \mathcal{C} -separable from L' . The \mathcal{C} -membership problem, which asks whether an input regular language belongs to \mathcal{C} , reduces to the \mathcal{C} -separation problem, as a regular language belongs to \mathcal{C} iff it is \mathcal{C} -separable from its complement.

Logics. We investigate several fragments of first-order logic on finite words. We view a finite word as a logical structure made of a sequence of positions labeled over A . We work with first-order logic $\text{FO}(<)$ using a unary predicate P_a for each $a \in A$, which selects positions labeled with an a , as well as binary predicates '=' for equality and '<' for the linear order. Such a formula defines the regular language of all words that satisfy it. We will freely use the name of a logical fragment of $\text{FO}(<)$ to denote the class of languages definable in this fragment. Observe that $\text{FO}(<)$ is powerful enough to express the following logical relations:

- First position, $\text{min}(x)$: $\forall y \neg(y < x)$.
- Last position, $\text{max}(x)$: $\forall y \neg(x < y)$.
- Successor, $y = x + 1$: $x < y \wedge \neg(\exists z x < z \wedge z < y)$.

However, for most fragments of $\text{FO}(<)$ this is not the case. For example, in the two-variables restriction $\text{FO}^2(<)$ of $\text{FO}(<)$, it is not possible to express successor, as it requires quantifying over a third variable. For these fragments \mathcal{F} , adding the predicates min, max

■ **Table 1** Logical fragments to which the technique applies.

Weak variant	FO(=)	FO ² (<)	$\Sigma_n(<)$	$\mathcal{B}\Sigma_n(<)$
Strong variant	FO(=, +1)	FO ² (<, +1)	$\Sigma_n(<, +1, \min, \max)$	$\mathcal{B}\Sigma_n(<, +1, \min, \max)$

and +1 yields a strictly more powerful logic \mathcal{F}^+ . Our goal is to prove a transfer result for such fragments: given a fragment, if the separation problem is decidable for the weak variant \mathcal{F} , then it is decidable as well for the strong variant \mathcal{F}^+ obtained by enriching \mathcal{F} with the above relations. The technique is *generic*, meaning that it is not bound to a particular logic. In particular, our transfer result applies to the following well-known logical fragments:

- FO(=), the restriction of FO(<) in which the linear order cannot be used, and only equality between two positions can be tested. The enriched fragment FO(=, +1) (*min* and *max* can be eliminated from the formulas) defines locally threshold testable languages [24].
- All levels in the quantifier alternation hierarchy of first-order logic. A first-order formula is $\Sigma_n(<)$ (resp. $\Pi_n(<)$) if its prenex normal form contains at most $(n - 1)$ quantifier alternations and starts with an \exists (resp. a \forall) quantifier block. Finally, a $\mathcal{B}\Sigma_n(<)$ formula is a boolean combination of $\Sigma_n(<)$ and $\Pi_n(<)$ formulas.
Since for all fragments above $\Sigma_2(<)$, a formula involving *min* and *max* can be expressed without these predicates in the same logic, we shall denote the enriched fragments by $\Sigma_1(<, +1, \min, \max)$, $\mathcal{B}\Sigma_1(<, +1, \min, \max)$, and then by $\Sigma_2(<, +1)$, $\mathcal{B}\Sigma_2(<, +1)$, ...
- FO²(<), the restriction of FO(<) using only two reusable variables. The corresponding enriched fragment is FO²(<, +1), since *min* and *max* can again be eliminated from the formulas.

Table 1 summarizes all fragments the technique applies to. We prove the following theorem.

► **Theorem 1.** *Let \mathcal{F} and \mathcal{F}^+ be respectively the weak and strong variants of one of the logical fragments in Table 1. Then \mathcal{F}^+ -separability can be effectively reduced to \mathcal{F} -separability.*

As explained in the introduction, we prove this theorem in two flavors: the first one, Theorem 4, is purely logical. It is self-contained and elementary, but is not entirely generic. The other one, Theorem 15, is purely algebraic and generic: the transfer works from an algebraic class (for which only fairly general restrictions are assumed) to an enriched one. Yet, it relies on already established results to be instantiated on the fragments of Table 1.

All these logical fragments have a rich history and have been extensively studied in the literature. In particular, the separation problem is known to be decidable for the following fragments: FO(=), FO²(<), $\Sigma_1(<)$, $\mathcal{B}\Sigma_1(<)$, $\Sigma_2(<)$ [4, 16, 17]. This means that, from our results, we obtain decidability of separation for FO(=, +1), FO²(<, +1), $\Sigma_1(<, +1, \min, \max)$, $\mathcal{B}\Sigma_1(<, +1, \min, \max)$ and $\Sigma_2(<, +1)$. Note that for FO(=, +1), FO²(<, +1) and $\mathcal{B}\Sigma_1(<, +1, \min, \max)$, the results could already be obtained as corollaries of algebraic theorems of Steinberg [21] and Almeida [2]. As explained in the introduction, an issue with this approach is that the proof of Steinberg's result relies on deep algebraic arguments and is not tailored to separation (the connection with separation is made by Almeida [2]). For $\Sigma_1(<, +1, \min, \max)$ and $\Sigma_2(<, +1)$, the result is new, as Steinberg's result does not apply to classes of languages that are not closed under complement.

3 Tools: Semigroups and Well-Formed Words

In this section, we define the main tools used in the paper. First, we recall the well-known semigroup based definition of regular languages: a language is regular if and only if it can be recognized by a finite semigroup. Our second tool, *well-formed words*, is specific to our problem and plays a key role in our transfer result.

3.1 Semigroups and Monoids

We work with the algebraic representation of regular languages in terms of semigroups. A *semigroup* is a set S equipped with an associative product, written $s \cdot t$ or st . A *monoid* is a semigroup S having a neutral element 1_S , *i.e.*, such that $s \cdot 1_S = 1_S \cdot s = s$ for all $s \in S$. If S is a semigroup, then S^1 denotes the monoid $S \cup \{1_S\}$ where $1_S \notin S$ is a new element, acting as neutral element. Note that we add such a new identity even if S is already a monoid.

An element $e \in S$ is *idempotent* if $e \cdot e = e$. We denote by $E(S)$ the set of idempotents of S . Given a *finite* semigroup S , it is folklore and easy to see that there is an integer $\omega(S)$ (denoted by ω when S is understood) such that for all s of S , s^ω is idempotent: $s^\omega = s^\omega s^\omega$.

Note that A^+ and A^* equipped with concatenation are respectively a semigroup and a monoid called the *free semigroup over A* and the *free monoid over A* . Let $L \subseteq A^+$ be a language and S be a semigroup (resp. a monoid). We say that L is *recognized by S* if there exist a morphism $\alpha : A^+ \rightarrow S$ (resp. $\alpha : A^* \rightarrow S$) and a set $F \subseteq S$ such that $L = \alpha^{-1}(F)$.

Semigroups and Separation. The separation problem takes as input two regular languages L, L' . It is convenient to work with a single object recognizing both of them, rather than having to deal with two. Let S, S' be semigroups recognizing L, L' together with the associated morphisms α, α' , respectively. Clearly, L and L' are both recognized by $S \times S'$ with the morphism $\alpha \times \alpha' : A^+ \rightarrow S \times S'$ mapping w to $(\alpha(w), \alpha'(w))$. From now on, we work with such a single semigroup recognizing both languages. Replacing $S \times S'$ with its image under $\alpha \times \alpha'$, one can also assume that this morphism is surjective. To sum up, we assume from now on, w.l.o.g., that L and L' are recognized by a single surjective morphism.

3.2 Well-Formed Words

In this section, we define our main tool for this paper. Assume that \mathcal{F} is the weak variant of one of the logical fragments of Table 1 and let \mathcal{F}^+ be the corresponding enriched variant. To any semigroup morphism $\alpha : A^+ \rightarrow S$ into a finite semigroup S , we associate a new alphabet \mathbb{A}_α called the alphabet of *well-formed words*. The main intuition behind this notion is that the \mathcal{F}^+ -separation problem for any two regular languages recognized by α can be reduced to the \mathcal{F} -separation problem for two regular languages over \mathbb{A}_α .

The alphabet \mathbb{A}_α , called *alphabet of well-formed words of α* , is defined from $\alpha : A^+ \rightarrow S$ by:

$$\mathbb{A}_\alpha = (E(S) \times S \times E(S)) \cup (S \times E(S)) \cup (E(S) \times S) \cup S.$$

We will not be interested in all words of \mathbb{A}_α^+ , but only in those that are well-formed. A word $w \in \mathbb{A}_\alpha^+$ is said to be *well-formed* if one of the following two properties holds:

- w is a single letter $s \in S$,
- w has length ≥ 2 and is of the form

$$(s_0, f_0) \cdot (e_1, s_1, f_1) \cdots (e_n, s_n, f_n) \cdot (e_{n+1}, s_{n+1}) \in (S \times E(S)) \cdot (E(S) \times S \times E(S))^* \cdot (E(S) \times S)$$

with $f_i = e_{i+1}$ for all $0 \leq i \leq n$.

► **Fact 2.** *The set of well-formed words of \mathbb{A}_α^+ is a regular language.*

We now define a morphism $\beta : \mathbb{A}_\alpha^+ \rightarrow S$ as follows. If $s \in S$, we set $\beta(s) = s$, if $(e, s) \in E(S) \times S$, we set $\beta((e, s)) = es$, if $(s, e) \in S \times E(S)$, we set $\beta((s, e)) = se$ and if $(e, s, f) \in E(S) \times S \times E(S)$, we set $\beta((e, s, f)) = esf$.

Associated Language of Well-formed Words. To any language $L \subseteq A^+$ that is recognized by α , one associates a language of well-formed words $\mathbb{L} \subseteq \mathbb{A}_\alpha^+$:

$$\mathbb{L} = \{w \in \mathbb{A}_\alpha^+ \mid w \text{ is well-formed and } \beta(w) \in \alpha(L)\}.$$

By definition, the language $\mathbb{L} \subseteq \mathbb{A}_\alpha^+$ is the intersection of the language of well-formed words with $\beta^{-1}(\alpha(L))$. Therefore, it is immediate by Fact 2 that it is regular, more precisely:

► **Fact 3.** *Let $L \subseteq A^+$ be recognized by α . Then, the associated language of well-formed words $\mathbb{L} \subseteq \mathbb{A}_\alpha^+$ is a regular language that one can effectively compute from a recognizer of L .*

4 Logical Approach

In this section, we prove Theorem 1 from a logical perspective. We begin with presenting our *separation* theorem, which will entail the *membership* theorem as a simple consequence.

► **Theorem 4.** *Let \mathcal{F} and \mathcal{F}^+ be respectively the weak and strong variants of one of the logical fragments in Table 1.*

Let L, L' be two languages recognized by a morphism $\alpha : A^+ \rightarrow S$ into a finite semigroup S . Let $\mathbb{L}, \mathbb{L}' \subseteq \mathbb{A}_\alpha^+$ be the languages of well-formed words associated with L, L' , respectively. Then L is \mathcal{F}^+ -separable from L' iff \mathbb{L} is \mathcal{F} -separable from \mathbb{L}' .

Theorem 4 reduces \mathcal{F}^+ -separation to \mathcal{F} -separation. The latter was already known to be decidable for several weak variants in Table 1, namely for $\text{FO}(=)$ [15], $\text{FO}^2(<)$ [16], $\Sigma_1(<)$ [4], $\mathcal{B}\Sigma_1(<)$ [4, 16] and $\Sigma_2(<)$ [17]. Hence, we get the following corollary.

► **Corollary 5.** *Let L, L' be regular languages. Then the following problems are decidable:*

- *whether L is $\text{FO}(=, +1)$ -separable from L' .*
- *whether L is $\text{FO}^2(<, +1)$ -separable from L' .*
- *whether L is $\Sigma_1(<, +1, \min, \max)$ -separable from L' .*
- *whether L is $\mathcal{B}\Sigma_1(<, +1, \min, \max)$ -separable from L' .*
- *whether L is $\Sigma_2(<, +1)$ -separable from L' .*

Notice that since the membership problem reduces to the separation problem, this also gives a new proof that all these fragments have a decidable membership problem. This is of particular interest for $\text{FO}^2(<, +1)$, $\mathcal{B}\Sigma_1(<, +1, \min, \max)$ and $\Sigma_2(<, +1)$ for which the previous proofs, which can be found in, or derived from [22, 1, 14], [8], and [6, 13, 12] respectively, are known to be quite involved. It turns out that for $\Sigma_2(<, +1)$, we can do even better and entirely avoid separation. Indeed, when \mathcal{F} is expressive enough, Theorem 4 can be used to prove a similar theorem for the membership problem.

► **Theorem 6.** *Let \mathcal{F} and \mathcal{F}^+ be respectively the weak and strong variants of one of the logical fragments in Table 1. Moreover, assume that for any alphabet of well-formed words, the set of well-formed words over this alphabet is definable in \mathcal{F} .*

Let L be a language recognized by a morphism $\alpha : A^+ \rightarrow S$ into a finite semigroup S . Let $\mathbb{L} \subseteq \mathbb{A}_\alpha^+$ be the language of well-formed words associated with L . Then L is definable in \mathcal{F}^+ iff \mathbb{L} is definable in \mathcal{F} .

Proof. Set $K = A^+ \setminus L$ and let \mathbb{K} be the associated language of well-formed words. Observe that by definition, $\mathbb{K} \cup \mathbb{L}$ is the set of all well-formed words.

If \mathbb{L} is definable in \mathcal{F} , then \mathbb{L} is \mathcal{F} -separable from \mathbb{K} , hence by Theorem 4, L is \mathcal{F}^+ -separable from K , and so L is definable in \mathcal{F}^+ . Conversely, if L is definable in \mathcal{F}^+ , then L is \mathcal{F}^+ -separable from K and by Theorem 4, \mathbb{L} is \mathcal{F} -separable from \mathbb{K} . Since $\mathbb{K} \cup \mathbb{L}$ is the set of all well-formed words, \mathbb{L} is the intersection of the separator with the set of all well-formed words, which by hypothesis is also definable in \mathcal{F} . Therefore, \mathbb{L} is definable in \mathcal{F} . ◀

Observe that being well-formed can be expressed in $\Pi_2(<)$: essentially, a word is well-formed if for all pairs of positions, either there is a third one in-between, or the labels of the two positions are “compatible”. Hence, among the fragments of Table 1, Theorem 6 applies to all fragments including and above $\Pi_2(<)$ in the quantifier alternation hierarchy. While such a transfer result was previously known [22, 13], the presentation and the proof are new. In particular, since membership is known to be decidable for $\Pi_2(<)$ [12], $\mathcal{B}\Sigma_2(<)$ [17] and $\Sigma_3(<)$ [17], we obtain new and simpler proofs of the following results.

► **Corollary 7.** *Given a regular language L , one can decide whether*

- *L is definable by a $\Sigma_2(<, +1)$ (resp. by a $\Pi_2(<, +1)$) formula.*
- *L is definable by a $\mathcal{B}\Sigma_2(<, +1)$ formula.*
- *L is definable by a $\Sigma_3(<, +1)$ (resp. by a $\Pi_3(<, +1)$) formula.*

It remains to prove Theorem 4. We devote the rest of the section to this proof. An important remark is that the proof of the right to left direction is constructive: we start with an \mathcal{F} formula that separates \mathbb{L} from \mathbb{L}' and use it to construct an \mathcal{F}^+ formula that separates L from L' . Note that the argument is generic for all fragments we consider.

On the other hand, the other direction, namely Proposition 9 below, requires a specific argument tailored to each fragment, which is a straightforward but tedious Ehrenfeucht-Fraïssé argument. Due to lack of space, we provide proofs of this proposition for each fragment in the long version [19] of this paper.

4.1 From \mathcal{F}^+ -separation to \mathcal{F} -separation

We prove that if L is \mathcal{F}^+ -separable from L' , then \mathbb{L} is \mathcal{F} -separable from \mathbb{L}' . We actually prove the contrapositive: if \mathbb{L} is *not* \mathcal{F} -separable from \mathbb{L}' , then L is *not* \mathcal{F}^+ -separable from L' . We rely on a construction which, to any well-formed word $\mathbf{u} \in \mathbb{A}_\alpha^+$ and any integer $i > 0$, associates a canonical word $\lceil \mathbf{u} \rceil_i \in A^+$.

Canonical Word Associated to a Well-formed Word. To any $s \in S$, we associate an arbitrarily chosen nonempty word $\lceil s \rceil \in A^+$ such that $\alpha(\lceil s \rceil) = s$ (which is possible since α has been chosen surjective). Let $i > 0$. From a well-formed word $\mathbf{u} \in \mathbb{A}_\alpha^+$, we build a word $\lceil \mathbf{u} \rceil_i \in A^+$ as follows. If $\mathbf{u} = s \in S$, then $\lceil \mathbf{u} \rceil_i = \lceil s \rceil$ for all i . Otherwise, we have by definition

$$\mathbf{u} = (s_0, e_1)(e_1, s_1, e_2) \cdots (e_{n-1}, s_{n-1}, e_n)(e_n, s_n).$$

For a natural $i > 0$, we set

$$\lceil \mathbf{u} \rceil_i = \lceil s_0 \rceil \lceil e_1 \rceil^i \lceil s_1 \rceil \lceil e_2 \rceil^i \cdots \lceil e_{n-1} \rceil^i \lceil s_{n-1} \rceil \lceil e_n \rceil^i \lceil s_n \rceil.$$

Recall that β is the morphism $\beta : \mathbb{A}_\alpha^+ \rightarrow S$ mapping \mathbf{u} to $s_0 e_1 s_1 \cdots s_{n-1} e_n s_n$. Since $e_j \in E(S)$ for all j , it is immediate that $\alpha(\lceil \mathbf{u} \rceil_i) = \beta(\mathbf{u})$, hence we get the following fact:

► **Fact 8.** For every $i > 0$ and every well-formed word $u \in \mathbb{A}_\alpha^+$, we have $u \in \mathbb{L}$ (resp. $u \in \mathbb{L}'$) if and only if $\lceil u \rceil_i \in L$ (resp. $u \in L'$).

We now proceed with the proof. We use the classical preorders associated to fragments of first-order logic. The (*quantifier*) *rank* of a first-order formula φ is the largest number of quantifiers along a branch in the parse tree of φ . Given $u, v \in A^+$, we write $u \preceq_k^{+1} v$ if any \mathcal{F}^+ formula of rank k that is satisfied by u is satisfied by v as well. Similarly, for $u, v \in \mathbb{A}_\alpha^+$, we write $u \preceq_k v$ if any \mathcal{F} formula of rank k that is satisfied by u is satisfied by v as well. One can verify that \preceq_k and \preceq_k^{+1} are preorders, as well as the following standard fact:

$$\begin{aligned} L \subset A^+ \text{ is definable by an } \mathcal{F}^+ \text{ formula of rank } k \text{ iff } L &= \{u' \mid \exists u \in L \text{ st. } u \preceq_k^{+1} u'\} \\ \mathbb{L} \subset \mathbb{A}_\alpha^+ \text{ is definable by an } \mathcal{F} \text{ formula of rank } k \text{ iff } \mathbb{L} &= \{u' \mid \exists u \in \mathbb{L} \text{ st. } u \preceq_k u'\}. \end{aligned} \quad (1)$$

Note that when \mathcal{F} and \mathcal{F}^+ are closed under complement, then \preceq_k and \preceq_k^{+1} are actually equivalence relations. We can now state the main proposition of this direction.

► **Proposition 9.** For any $k \in \mathbb{N}$, there exist $\ell \in \mathbb{N}$ and $i \in \mathbb{N}$ such that for any well-formed words $u, u' \in \mathbb{A}_\alpha^+$ satisfying $u \preceq_\ell u'$, we have $\lceil u \rceil_i \preceq_k^{+1} \lceil u' \rceil_i$.

For all fragments of Table 1, Proposition 9 is proved using classical Ehrenfeucht-Fraïssé arguments. While each proof is specific, the underlying ideas are similar. We present these proofs in the long version of this paper [19]. We finish the subsection by explaining how Proposition 9 can be used to terminate the proof of the first direction of Theorem 4.

We argue by contrapositive: assume that \mathbb{L} is *not* \mathcal{F} -separable from \mathbb{L}' . By definition this means that no language definable in \mathcal{F} separates \mathbb{L} from \mathbb{L}' . In particular, for any ℓ , the language

$$\{u' \mid \exists u \in \mathbb{L} \text{ st. } u \preceq_\ell u'\},$$

which is definable in \mathcal{F} by (1), cannot be a separator. Note that this language contains \mathbb{L} . Hence, for all $\ell \in \mathbb{N}$, there exist $u \in \mathbb{L}$ and $u' \in \mathbb{L}'$ such that $u \preceq_\ell u'$. We deduce from Proposition 9 and Fact 8 that for all $k \in \mathbb{N}$, there exist $u \in L$ and $u' \in L'$ such that $u \preceq_k^{+1} u'$. It follows, again by (1), that L is *not* \mathcal{F}^+ -separable from L' , which terminates the proof.

4.2 From \mathcal{F} -separation to \mathcal{F}^+ -separation

We now prove that if \mathbb{L} is \mathcal{F} -separable from \mathbb{L}' , then L is \mathcal{F}^+ -separable from L' . We do so by building an \mathcal{F}^+ -definable separator. This proof is this time entirely generic. We rely on a construction that is dual to the one used previously: to any word $w \in A^+$, we associate a canonical well-formed word $\lfloor w \rfloor \in \mathbb{A}_\alpha^+$.

Canonical Well-formed Word Associated to a Word. To any word w of A^+ , we associate a canonical well-formed word $\lfloor w \rfloor \in \mathbb{A}_\alpha^+$ such that $\alpha(w) = \beta(\lfloor w \rfloor)$. This construction is adapted from [14] and is originally inspired by [22].

Fix an arbitrary order on the set $E(S)$. For a position x of w , let $u_x \in A^+$ be the infix of w obtained by keeping only positions $x - (|S| - 1)$ to x . If position $x - (|S| - 1)$ does not exist, u_x is just the prefix of w ending at x . A position x is said *distinguished* if there exists an idempotent $e \in E(S)$ such that $\alpha(u_x) \cdot e = \alpha(u_x)$. Additionally, we always define the rightmost position as distinguished, even if it does not satisfy the property. Set $x_1 < \dots < x_{n+1}$ as the distinguished positions in w , so that x_{n+1} is the rightmost position. Let $e_1, \dots, e_n \in E(S)$ be such that for all $1 \leq i \leq n - 1$, e_i is the smallest idempotent such that $\alpha(u_{x_i}) \cdot e_i = \alpha(u_{x_i})$.

If $n = 0$, *i.e.*, if the only distinguished position is the rightmost one, set $\lfloor w \rfloor = \alpha(w) \in \mathbb{A}_\alpha$. Otherwise, we define $\lfloor w \rfloor \in \mathbb{A}_\alpha^+$ as the word:

$$\lfloor w \rfloor = (\alpha(w_0), e_1) \cdot (e_1, \alpha(w_1), e_2) \cdots (e_{n-1}, \alpha(w_{n-1}), e_n) \cdot (e_n, \alpha(w_n)) \tag{2}$$

where w_0 is the prefix of w ending at position x_1 , for all $1 \leq i \leq n - 1$, w_i is the infix of w obtained by keeping positions $x_i + 1$ to x_{i+1} , and w_n is the suffix of w starting at position $x_n + 1$. Note that by construction, $\lfloor w \rfloor$ is well-formed.

The next statement follows from the definition of β , and from the fact that by definition of the words w_i and of the chosen idempotents, we have $\alpha(w_0 \cdots w_i) e_{i+1} = \alpha(w_0 \cdots w_i)$.

► **Fact 10.** *For all $w \in A^+$, we have $\alpha(w) = \beta(\lfloor w \rfloor)$. Therefore, $w \in L$ iff $\lfloor w \rfloor \in \mathbb{L}$ and $w \in L'$ iff $\lfloor w \rfloor \in \mathbb{L}'$.*

To any distinguished position x_i in w , we now associate the position $\lfloor x \rfloor = i$ in $\lfloor w \rfloor$. Our main motivation for using this construction is its local canonicity, which is stated in the following lemma.

► **Lemma 11.** *Let $w \in A^+$. Then we have the following properties:*

- (a) *whether a position x is distinguished in w , and if so the label of position $\lfloor x \rfloor$ in $\lfloor w \rfloor$ only depends on the infix of w of length $2|S|$ ending at position x . That is, if the infixes of length $2|S|$ ending at x and y are equal, then x is distinguished iff so is y , and in that case, the labels of $\lfloor x \rfloor$ and $\lfloor y \rfloor$ in $\lfloor w \rfloor$ are equal.*
- (b) *the label of the last position of $\lfloor w \rfloor$ only depends on the suffix of length $2|S|$ of w .*

Proof. It is immediate that whether x is distinguished and if so the associated idempotent only depends on the infix u_x of length at most $|S|$ ending at x . Therefore, to prove (a), it suffices to show that all infixes w_i used in (2) are of size at most $|S|$, or in other words, that among $|S| + 1$ consecutive positions, at least one is distinguished. So let us consider an infix $a_1 \cdots a_{|S|+1}$ of w of length $|S| + 1$. It is immediate from the pigeonhole principle that there exist $i < j$ such that $\alpha(a_1 \cdots a_i) = \alpha(a_1 \cdots a_j) = \alpha(a_1 \cdots a_i) \cdot (\alpha(a_{i+1} \cdots a_j))^\omega$. Hence, the position corresponding to a_i is distinguished. The proof of the second assertion is similar. ◀

L is \mathcal{F}^+ -separable from L' . We can now construct our separator. The construction follows from the next proposition.

► **Proposition 12.** *Let $\mathbb{K} \subseteq \mathbb{A}_\alpha^+$ that can be defined using an \mathcal{F} formula φ . Then there exists an \mathcal{F}^+ formula Ψ over alphabet A such that for every word $w \in A^+$:*

$$w \models \Psi \text{ if and only if } \lfloor w \rfloor \models \varphi.$$

Proof. Proposition 12 follows from the following simple consequence of Lemma 11.

► **Claim 13.** *For any $\mathfrak{a} \in \mathbb{A}_\alpha$ there exists a formula $\gamma_{\mathfrak{a}}(x)$ of \mathcal{F}^+ with a free variable x , such that for any $w \in A^+$ and any position x of w , we have $w \models \gamma_{\mathfrak{a}}(x)$ iff x is distinguished and $\lfloor x \rfloor$ has label \mathfrak{a} in $\lfloor w \rfloor$.*

This claim holds since by Lemma 11, formula $\gamma_{\mathfrak{a}}(x)$ only needs to explore the neighborhood of size $2|S|$ of x , which is trivially possible for all fragments \mathcal{F}^+ we consider. To conclude the proof of Proposition 12, it suffices to define Ψ as the formula constructed from φ by restricting all quantifiers to positions that are distinguished and to replace all tests $P_{\mathfrak{a}}(x)$ by $\gamma_{\mathfrak{a}}(x)$. ◀

We can now finish the proof of Theorem 4. Assume that \mathbb{L} is \mathcal{F} -separable from \mathbb{L}' and let φ be an \mathcal{F} formula defining a separator. We denote by Ψ the \mathcal{F}^+ formula obtained from φ as defined in Proposition 12. We prove that Ψ defines a language separating L from L' .

We first prove that $L \subseteq \{w \mid w \models \Psi\}$. Assume that $w \in L$. Then by Fact 10, we have $\lfloor w \rfloor \in \mathbb{L}$. Hence, $\lfloor w \rfloor \models \varphi$ and so $w \models \Psi$ by definition of Ψ . The proof that $L' \subseteq \{w \mid w \not\models \Psi\}$ is identical: if $w \in L'$, we have $\lfloor w \rfloor \in \mathbb{L}'$ by Fact 10. Hence, $\lfloor w \rfloor \not\models \varphi$ and $w \not\models \Psi$ by definition of Ψ . \blacktriangleleft

5 Algebraic Approach

We now present an algebraic version of Theorem 4: the operator $\mathbb{V} \mapsto \mathbb{V} * \mathbb{D}$ preserves decidability of separation.

We would like to emphasize again that the ideas behind this theorem are essentially the same as for Theorem 4. In particular, proofs presented in the long version of this paper [19] only rely on elementary notions, thus bypassing complex constructions usually used to prove this kind of result, even if the statement itself requires some additional algebraic vocabulary.

The section is organized in three parts.

- We first briefly recall how classes of languages corresponding to our logical fragments are given an algebraic definition: for each fragment, an associated class of finite semigroups (or monoids) \mathbb{V} , a *variety*, has already been characterized, such that the class of languages definable in the fragment is exactly the class of languages that are recognized by a semigroup (or monoid) of \mathbb{V} .
- In the second part, we define what “adding the successor relation” means in this context. Given a variety \mathbb{V} , this generally corresponds to considering a new variety built on top of \mathbb{V} via an operation called the *semidirect product*. This new variety is denoted $\mathbb{V} * \mathbb{D}$.
- Finally, in the last part, we state our main theorem: for any variety \mathbb{V} , separability for the variety $\mathbb{V} * \mathbb{D}$ reduces to separability for the variety \mathbb{V} .

5.1 Varieties

A *variety of semigroups (resp. monoids)* is a class of finite semigroups (resp. monoids) closed under three natural operations: finite direct product, subsemigroup (or submonoid), and homomorphic image. A variety \mathbb{V} defines a class of languages, also noted \mathbb{V} , namely the class of all of languages recognized by semigroups (resp. monoids) in \mathbb{V} . There is an issue however: all classes of languages defined in this way have to be closed under complement, since the set of languages recognized by any semigroup is closed under complement. This prevents us from capturing logical fragments that are not closed under complement, such as $\Sigma_2(<)$. This problem has been solved in [11] with the notions of *ordered semigroups and monoids*. Intuitively, such a semigroup is parametrized by a partial order and the set of languages it recognizes is then restricted with respect to this partial order. These classical constructions will be recalled in the long version of this paper [19], as well as varieties corresponding to all fragments we deal with.

All logical fragments presented in Section 2 correspond to varieties that have been fully identified. For each fragment, its non-enriched variant corresponds to a variety \mathbb{V} of (ordered) monoids and its enriched version to the variety of (ordered) semigroups $\mathbb{V} * \mathbb{D}$ built from \mathbb{V} . For example, the fragment $\text{FO}^2(<)$ corresponds to the variety of monoids DA and the fragment $\text{FO}^2(<, +1)$ to the variety of semigroups $\text{DA} * \mathbb{D}$ [23] (see the long version [19] for a bibliography with all correspondences).

5.2 Semidirect Product

The Variety \mathbf{D} . The variety \mathbf{D} consists of all finite ordered semigroups S such that for all $s \in S$ and all $e \in E(S)$, we have $se = e$. From a language perspective, a language L is recognized by a semigroup in \mathbf{D} iff there exists $k \in \mathbb{N}$ such that membership of a word w to L only depends on the suffix of length k of w .

Semidirect Product. Let M be an ordered monoid and let T be an ordered semigroup. A *semidirect product* of M and T is an operation that is parametrized by an *action* of T on M and outputs a new ordered semigroup, whose base set is $M \times T$. Therefore, one can obtain different semidirect products out of the same M and T , depending on the chosen action (we recall the construction in the long version [19]). One can next lift this product at the level of varieties.

We are interested in the semidirect products of the form $\mathbf{V} * \mathbf{D}$, the variety of ordered semigroups generated by all semidirect products of an ordered monoid of \mathbf{V} by an ordered semigroup of \mathbf{D} . The reason why we introduce such semidirect products is the following theorem, which gathers several nontrivial results from the literature. The reader is referred to the long version of this paper [19] for details.

► **Theorem 14.** *Let \mathbf{V} be a variety corresponding to a fragment \mathcal{F} from the ones presented in Table 1. Then, the variety corresponding to the fragment \mathcal{F}^+ is $\mathbf{V} * \mathbf{D}$.*

5.3 Main Theorem

We have now the machinery needed to state our main theorem. For any variety of ordered monoids \mathbf{V} , we reduce $(\mathbf{V} * \mathbf{D})$ -separability to \mathbf{V} -separability.

► **Theorem 15.** *Let \mathbf{V} be a non-trivial variety of ordered monoids. Let L and L' be two languages both recognized by the same morphism $\alpha : A^+ \rightarrow S$ into a finite semigroup S . Set $\mathbb{L}, \mathbb{L}' \subseteq \mathbb{A}_\alpha^+$ as the languages of well-formed words associated to L, L' , respectively. Then, L is $(\mathbf{V} * \mathbf{D})$ -separable from L' if and only if \mathbb{L} is \mathbf{V} -separable from \mathbb{L}' .*

The proof of Theorem 15 is presented in the full version of this paper [19]. As it was the case for Theorem 4, the proof is both elementary and constructive: if there exists a separator for \mathbb{L} and \mathbb{L}' in \mathbf{V} , we use it to construct a separator for L and L' in $\mathbf{V} * \mathbf{D}$.

In view of Theorem 14, Theorem 15 applies to all fragments we introduced. This means that Theorem 4 can be given an alternate indirect proof within this algebraic framework by combining Theorem 15 and Theorem 14. Hence, this also yields another proof of Corollary 5.

6 Conclusion

We proved that separation is decidable over finite words for the following logical fragments: $\text{FO}(=, +1)$, $\Sigma_1(<, +1, \min, \max)$, $\mathcal{B}\Sigma_1(<, +1, \min, \max)$, $\Sigma_2(<, +1)$ and $\text{FO}^2(<, +1)$. To achieve this, we presented a simple reduction to the same problem for the weaker fragments $\text{FO}(=)$, $\Sigma_1(<)$, $\mathcal{B}\Sigma_1(<)$, $\Sigma_2(<)$ and $\text{FO}^2(<)$.

The reduction itself is entirely generic to all fragments and its proof is elementary, and also mostly generic. In particular, the technique can be used to prove that the reduction works for other natural fragments of first-order logic. An interesting example to which these results apply is the quantifier alternation hierarchy within $\text{FO}^2(<)$ (known as the Trotter-Weil hierarchy, and which is decidable [25]). However, the separation problem for

classes in this hierarchy has yet to be investigated. We also obtained direct proofs that membership is decidable for $\mathcal{B}\Sigma_2(<, +1)$ and $\Sigma_3(<, +1)$.

Finally, we presented an algebraic formulation of this reduction, which recovers a previously known result by Steinberg [21], while having a much simpler proof. One can expect extending these results to other fragments, such as enrichment with modulo predicates. Another advantage of this technique is that it can be extended in a straightforward way to the same logical fragments over words of infinite length. This yields identical transfer results. We leave the presentation of these results for further work.

References

- 1 Jorge Almeida. A syntactical proof of locality of DA. *International Journal on Algebra and Computation*, 6:165–177, 1996.
- 2 Jorge Almeida. Some algorithmic problems for pseudovarieties. *Publicationes Mathematicae Debrecen*, 54:531–552, 1999. Proc. of Automata and Formal Languages, VIII.
- 3 Karl Auinger. On the decidability of membership in the global of a monoid pseudovariety. *International Journal on Algebra and Computation*, 20(2):181–188, 2010.
- 4 Wojciech Czerwiński, Wim Martens, and Tomáš Masopust. Efficient separability of regular languages by subsequences and suffixes. In *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming, ICALP’13*, volume 7966 of *Lecture Notes in Computer Science*, pages 150–161. Springer, 2013.
- 5 Volker Diekert and Paul Gastin. First-order definable languages. In *Logic and Automata: History and Perspectives*, volume 2, pages 261–306. Amsterdam University Press, 2008.
- 6 Christian Glaßer and Heinz Schmitz. Languages of dot-depth $3/2$. *Theory of Computing Systems*, 42(2):256–286, 2008.
- 7 Karsten Henckell. Pointlike sets: the finest aperiodic cover of a finite semigroup. *Journal of Pure and Applied Algebra*, 55(1-2):85–126, 1988.
- 8 Robert Knast. A semigroup characterization of dot-depth one languages. *Rairo Informatique Théorique et Applications*, 17(4):321–330, 1983.
- 9 Manfred Kufleitner and Alexander Lauser. Around dot-depth 1. *International Journal of Foundations of Computer Science*, 23(6):1323–1340, 2012.
- 10 Robert McNaughton and Seymour Papert. *Counter-Free Automata*. MIT Press, 1971.
- 11 Jean-Éric Pin. A variety theorem without complementation. *Russian Mathematics, (Izvestija vuzov. Matematika)*, 39:80–90, 1995.
- 12 Jean-Éric Pin and Pascal Weil. Polynomial closure and unambiguous product. *Theory of Computing Systems*, 30(4):383–422, 1997.
- 13 Jean-Éric Pin and Pascal Weil. The wreath product principle for ordered semigroups. *Communications in Algebra*, 30:5677–5713, 2002.
- 14 Thomas Place and Luc Segoufin. Decidable characterization of $\text{FO}^2(<, +1)$ and locality of DA. Unpublished, to appear, 2014.
- 15 Thomas Place, Larijn van Rooijen, and Marc Zeitoun. Separating regular languages by locally testable and locally threshold testable languages. In *Proceedings of the 34th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS’13*, volume 24 of *LIPICs*, pages 363–375. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2013.
- 16 Thomas Place, Larijn van Rooijen, and Marc Zeitoun. Separating regular languages by piecewise testable and unambiguous languages. In *Proceedings of the 28th MFCS’13*, volume 8087 of *Lecture Notes in Computer Science*, pages 729–740. Springer, 2013.
- 17 Thomas Place and Marc Zeitoun. Going higher in the first-order quantifier alternation hierarchy on words. In *Proceedings of the 41th International Colloquium on Automata,*

- Languages, and Programming, ICALP'14*, volume 8573 of *Lecture Notes in Computer Science*, pages 342–353, 2014. <http://arxiv.org/pdf/1404.6832v1>.
- 18 Thomas Place and Marc Zeitoun. Separating regular languages with first-order logic. In *Proceedings of the Joint Meeting of the 23rd EACSL Annual Conference on Computer Science Logic (CSL'14) and the 29th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS'14)*, 2014.
 - 19 Thomas Place and Marc Zeitoun. A transfer theorem for the separation problem. *CoRR*, abs/1501.00569, 2015. <http://arxiv.org/abs/1501.00569>.
 - 20 Marcel-Paul Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8:190–194, 1965.
 - 21 Benjamin Steinberg. A delay theorem for pointlikes. *Semigroup Forum*, 63(3):281–304, 2001.
 - 22 Howard Straubing. Finite semigroup varieties of the form $V * D$. *Journal of Pure and Applied Algebra*, 36:53–94, 1985.
 - 23 Denis Thérien and Thomas Wilke. Over words, two variables are as powerful as one quantifier alternation. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing, STOC'98*, pages 234–240. ACM, 1998.
 - 24 Wolfgang Thomas. Classifying regular events in symbolic logic. *Journal of Computer and System Sciences*, 25(3):360–376, 1982.
 - 25 Manfred Kufleitner Pascal Weil. On logical hierarchies within FO^2 -definable languages. *Logical Methods in Computer Science*, 8(3), 2012.