An effective characterization of the alternation hierarchy in two-variable logic

Andreas Krebs[1] and Howard Straubing[2]

1 Wilhelm-Schickard-Institut, Universität Tübingen
Sand 13, 72076 Tübingen, Germany
mail@krebs-net.de*

2 Computer Science Department, Boston College
Chestnut Hill, Massachusetts, USA 02467
straubin@cs.bc.edu†

Abstract

We characterize the languages in the individual levels of the quantifier alternation hierarchy of first-order logic with two variables by identities. This implies decidability of the individual levels. More generally we show that two-sided semidirect products with **J** as the right-hand factor preserve decidability.

1998 ACM Subject Classification F.4.3 Formal Languages (D.3.1)

Keywords and phrases FO_2, Quantifier Alternation, J, Pseudovarities, Identities

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2012.86

1 Introduction

It has been known for some time (Kamp [6], Immerman and Kozen [5]) that every first-order sentence over the base $<$ defining properties of finite words is equivalent to one containing only three variables. The fragment $\mathrm{FO}^2[<]$ of sentences that use only two variables, has been the object of intensive study; Tesson and Thérien [18] give a broad-ranging survey of the many places in which the class of languages definable in this logic arises. Weis and Immerman [21] initiated the study of the hierarchy within $\mathrm{FO}^2[<]$ based on alternation of quantifiers. They showed, using model-theoretic techniques, that the hierarchy is infinite, but finite for each fixed alphabet.

In [17], the second author provided an algebraic characterization of the levels of the hierarchy, showing that they correspond to the levels of weakly iterated two-sided semidirect products of the pseudovariety **J** of finite $\mathcal{J}$-trivial monoids. This still left open the problem of *decidability* of the hierarchy: effectively determining from a description of a regular language the lowest level of the hierarchy to which the language belongs. This problem was apparently solved in Almeida-Weil [2], from which explicit identities for the iterated product varieties can be extracted. However, an error in that paper called the correctness of these results into question. Here we show that the given identities do indeed characterize these pseudovarieties. In particular, since it is possible to verify effectively whether a given finite monoid satisfies one of these identities, we obtain an effective procedure for exactly determining the alternation depth of a regular language definable in two-variable logic.

* Research partially supported by DFG/KR 4240/1 “Language Recognition by Algebra and Topology”
† Research partially supported by National Science Foundation Grant CCF-0915065

32nd Int’l Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2012).
Editors: D. D’Souza, J. Radhakrishnan, and K. Telikepalli; pp. 86–98
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

We show more generally that the two-sided semidirect product of a pseudovariety with $\mathbf{J}$ as the right-hand factor preserves decidability. That is, if we have an effective procedure for determining if a given finite monoid belongs to a pseudovariety $\mathbf{V}$, then we have such a procedure for $\mathbf{V} ** \mathbf{J}$.

At several junctures, our proof could have been shortened by appealing to known results about the algebra of finite categories and the topological theory of profinite monoids, which are the principal tools of [2]. For example, Theorem 5 is really just the bonded component theorem of Tilson [20] coupled with Simon's Theorem [15] on $\mathcal{J}$-trivial monoids. Lemma 7 closely mirrors the work of Almeida on the structure of the free profinite $\mathcal{J}$-trivial monoid [1]. In order to keep our argument accessible and self-contained, we have chosen to steer clear of these quite technical results. We do discuss finite categories, but only at the most elementary level. Avoiding profinite techniques forces us to give explicit size bounds, but these are of independent interest in decidability questions.

We give the necessary preliminaries from algebra in Section 2. Section 3 is devoted to our fundamental theorem, a category-based characterization of two-sided semidirect products with $\mathbf{J}$ as the right-hand factor. We apply this result in Section 4 to obtain explicit identities for the levels of the hierarchy, thus solving the decidability problem. We use these identities in Section 5 to give a new proof of the result of Weis and Immerman that the hierarchy collapses for each fixed input alphabet. Section 6 proves the general decidability-preserving result for block products with $\mathbf{J}$.

After we circulated an early draft of this paper, we became aware of a number of related results. Kufleitner and Weil [9], building on earlier work of theirs [8], independently established the decidability of the levels of the alternation hierarchy, using an entirely different algebraic characterization. A proof that $\mathbf{V} ** \mathbf{J}$ is decidable if $\mathbf{V}$ is appears in the unpublished Ph.D. thesis of Steinberg [16].

2 Preliminaries

While the principal application of our results is in finite model theory, this paper contains no formal logic *per se* and is entirely algebraic in content. The reader should consult [17] and [21] for the definition of $\mathrm{FO}^2[<]$ and the alternation hierarchy within it. For our purposes here, they are to be viewed simply as the language classes corresponding to certain varieties of finite monoids, as discussed below.

2.1 Finite monoids and regular languages

See the book by Pin [11] for a detailed treatment of the matters discussed in this subsection and the next; here we give a brief informal review.

A *monoid* is a set M together with an associative operation for which there is an identity element $1 \in M$. If A is a finite alphabet, then A^* is a monoid with concatenation of words as the multiplication. A^* is the *free monoid* on A: this means that every map $\alpha : A \to M$, where M is a monoid, extends in a unique fashion to a homomorphism from A^* into M.

Apart from free monoids, all the monoids we consider in this paper are finite. If M is a finite monoid, then for every element $m \in M$ there is a unique $e \in \{m^k : k > 1\}$ that is *idempotent, i.e.,* $e^2 = e$. We denote this element m^ω.

If M, N are monoids then we say M *divides* N, and write $M \prec N$, if M is a homomorphic image of a submonoid of N.

We are interested in monoids because of their connection with automata and regular languages: A *congruence* on A^* is an equivalence relation $\sim$ on A^* such that $u_1 \sim u_2$,

$v_1 \sim v_2$, implies $u_1v_1 \sim u_2v_2$. The classes of $\sim$ then form a monoid $M = A^*/\sim$, and the map $u \mapsto [u]_\sim$ sending each word to its congruence class is a homomorphism. If $L \subseteq A^*$, then $\equiv_L$, the *syntactic congruence* of L, is the coarsest congruence for which L is a union of congruence classes. The quotient monoid $A^*/\equiv_L$ is called the *syntactic monoid* of L and is denoted $M(L)$.

We say that a monoid M *recognizes* a language $L \subseteq A^*$ if there is a homomorphism $\alpha : A^* \to M$ and a subset X of M such that $\alpha^{-1}(X) = L$. The following proposition gives the fundamental properties linking automata to finite monoids.

▶ **Proposition 1.** *A language $L \subseteq A^*$ is regular if and only if $M(L)$ is finite. A monoid M recognizes L if and only if $M(L) \prec M$.*

2.2 Varieties and identities

A collection $\mathbf{V}$ of finite monoids closed under finite direct products and division is called a *pseudovariety* of finite monoids. (The prefix 'pseudo' is there because of the restriction to finite products, as the standard use of 'variety' in universal algebra does not carry this restriction.)

Given a pseudovariety $\mathbf{V}$, we consider for each finite alphabet A the set $A^*\mathcal{V}$ of regular languages $L \subseteq A^*$ such that $M(L) \in \mathbf{V}$. We call $\mathcal{V}$ the *variety of languages* corresponding to the pseudovariety $\mathbf{V}$. The correspondence $\mathbf{V} \mapsto \mathcal{V}$ is one-to-one, a consequence of the fact that every pseudovariety is generated by the syntactic monoids it contains. We are interested in this correspondence because of its connection with decidability problems for classes of regular languages: To test whether a given language L belongs to $A^*\mathcal{V}$, we compute its syntactic monoid $M(L)$ and test whether $M(L) \in \mathbf{V}$. Since the multiplication table of the syntactic monoid can be effectively computed from any automaton representation of L, decidability for the classes $A^*\mathcal{V}$ reduces to determining whether a given finite monoid belongs to $\mathbf{V}$.

Let Ξ be the countable alphabet $\{x_1, x_2, \ldots\}$. A *term* over Ξ is built from the letters by concatenation and application of a unary operation $v \mapsto v^\omega$. For example, $(x_1x_2)^\omega x_1$ is a term. We will interpret these terms in finite monoids in the obvious way, by considering a valuation $\psi : \Xi \to M$ and extending it to terms by giving concatenation and the ω operator their usual meaning in M. For this reason, we do not distinguish between $(uv)w$ and $u(vw)$, where u, v and w are themselves terms, nor between terms u^ω and $(u^\omega)^\omega$, as these will be equivalent under every valuation.

An *identity* is a formal equation $u = v$, where u and v are terms. We say that a monoid M *satisfies* the identity, and write $M \models (u = v)$, if u and v are equal under every valuation into M. The family of all finite monoids satisfying a given set of identities is a pseudovariety, and we say that the pseudovariety is *defined* by the set of identities. We must stress that the identities we consider here are very special instances of a much more general class of *pseudoidentities*. Under this broader definition, every pseudovariety is defined by a set of pseudoidentities. See, for instance, Almeida [1]. If a pseudovariety $\mathbf{V}$ is defined by a *finite* set of identities of the form we described, then membership of a given finite monoid M in $\mathbf{V}$ is decidable, since we only need to substitute elements of $\mathbf{V}$ for the variables in the identities in every way possible, and check that equality holds in each case.

We consider four particular pseudovarieties that will be of importance in this paper. (In presenting identities we will relax the formal requirement that all terms are over the alphabet $\{x_1, x_2, \ldots\}$, and use a larger assortment of letters for the variables.)

Ap The pseudovariety **Ap** consists of the *aperiodic* finite monoids, those that contain no nontrivial groups. It is defined by the identity $x^\omega = xx^\omega$. If A is a finite alphabet, then $M(L) \in \mathbf{Ap}$ if and only if L is definable by a first-order sentence over $<$. In other words, the first-order definable languages form the variety of languages corresponding to $\mathbf{Ap}$ (McNaughton and Papert [10]).

DA The pseudovariety **DA** is defined by the identity

$$(xyz)^\omega y(xyz)^\omega = (xyz)^\omega.$$

There are many equivalent characterizations of this pseudovariety in terms of other identities, the ideal structure of the monoids, and logic. For us the most important ones are these: First, **DA** is also defined by the identities ([14])

$$(xy)^\omega (yx)^\omega (xy)^\omega = (xy)^\omega,\ x^\omega = xx^\omega.$$

Second, let $e \in M$ be idempotent, and let M_e be the submonoid of M generated by the elements $m \in M$ for which $e \in MmM$. Then $M \in \mathbf{DA}$ if and only if $e = eM_e e$ for all idempotents e of M. Finally, if $L \subseteq A^*$ is a language, then $M(L) \in \mathbf{DA}$ if and only if L is definable in $\mathrm{FO}^2[<]$. In other words, the two-variable definable languages form the variety of languages corresponding to **DA**(Thérien and Wilke, [19]).

J The pseudovariety **J** consists of finite monoids that satisfy the pair of identities

$$(xy)^\omega = (yx)^\omega,\ x^\omega = xx^\omega.$$

This is equivalent to the identities

$$(xy)^\omega x = y(xy)^\omega = (xy)^\omega.$$

Alternatively, **J** consists of finite monoids M such that for all $s, t \in M$, $MsM = MtM$ implies $s = t$. Such monoids are said to be $\mathcal{J}$-trivial.

A theorem due to Imre Simon [15] describes the regular languages whose syntactic monoids are in **J**. Let $w \in A^*$. We say that $v = a_1 \cdots a_k$, where each $a_i \in A$, is a *subword* of w if $w = w_0 a_1 w_1 \cdots a_k w_k$ for some $w_i \in A^*$. We define an equivalence relation $\sim_k$ on A^* that identifies two words if and only if they contain the same subwords of length no more than k. In particular, $w_1 \sim_1 w_2$ if and only if w_1 and w_2 contain the same set of letters. Simon's theorem is:

▶ **Theorem 2.** *Let $\phi : A^* \to M$ be a homomorphism onto a finite monoid. Then the following are equivalent:*

- *$M \in \mathbf{J}$.*
- *There exists $k \geq 1$ such that if $w \sim_k w'$, then $\phi(w) = \phi(w')$. (In particular, M is a quotient of $A^* / \sim_k$.)*

It is easy to show that the second condition implies the first; the deep content of the theorem is the converse implication. The theorem can also be formulated in first-order logic: The variety of languages corresponding to **J** consists of languages definable by Boolean combinations of Σ_1 sentences over $<$.

$\mathbf{J_1}$ The pseudovariety $\mathbf{J_1}$ consists of all idempotent and commutative monoids; *i.e.,* those finite monoids that satisfy the identities $x^2 = x,\ xy = yx$. A language $L \subseteq A^*$ is in the variety of languages corresponding to $\mathbf{J_1}$ if and only if it is a union of $\sim_1$-classes. It is well known, and easy to show, that $\mathbf{J_1} \subseteq \mathbf{J} \subseteq \mathbf{DA} \subseteq \mathbf{Ap}$, and all the inclusions are proper.

2.3 Two-sided Semidirect Products

In this section we describe an operation on pseudovarieties of finite monoids, the *two-sided semidirect product.* This was given in its formal description by Rhodes and Tilson [12], but it has precursors in automata theory in the work of Schützenberger on sequential bimachines [13], Krohn, Mateosian and Rhodes [7], and Eilenberg on triple products [4]. Traditionally, one begins with a two-sided semidirect product operation on monoids, and then uses this to define the corresponding operation on pseudovarieties. Here we find it simpler to define the operation on varieties directly.

Let A be a finite alphabet, and $\psi : A^* \to N$ a homomorphism into a finite monoid. Let $\Sigma = N \times A \times N$, which we treat as a new finite alphabet. We define a length-preserving transduction (not a homomorphism) $\tau_\psi : A^* \to \Sigma^*$ by $\tau_\psi(1) = 1$, and

$$\tau_\psi(a_1 \cdots a_n) = \sigma_1 \cdots \sigma_n, \text{ where}$$

$$\sigma_i = (\psi(a_1 \cdots a_{i-1}), a_i, \psi(a_{i+1} \cdots a_n)) \in \Sigma.$$

(If $i = 1$, we interpret the right-hand side as $(1, a_1, \psi(a_2 \cdots a_n))$, and similarly if $i = n$.)

Let $\mathbf{V}$ and $\mathbf{W}$ be pseudovarieties of finite monoids. Let M be a finite monoid, and let $\phi : A^* \to M$ be a surjective homomorphism. We say that $M \in \mathbf{V} ** \mathbf{W}$ if and only if there exist homomorphisms

$$\psi : A^* \to N \in \mathbf{W},$$

$$h : (N \times A \times N)^* \to K \in \mathbf{V},$$

such that ϕ factors through $(h \circ \tau_\psi, \psi)$—in other words, for all $v, w \in A^*$, if $\psi(v) = \psi(w)$ and $h(\tau_\psi(v)) = h(\tau_\psi(w))$, then $\phi(v) = \phi(w)$. It is not difficult to check that this is independent of the alphabet A and the homomorphism ϕ, and is thus determined entirely by M, and that furthermore $\mathbf{V} ** \mathbf{W}$ forms a pseudovariety of finite monoids. We will treat this as the definition of $\mathbf{V} ** \mathbf{W}$, but it is also straightforward to verify that this coincides with the pseudovariety generated by two-sided semidirect products $K ** N$, where $K \in \mathbf{V}$ and $N \in \mathbf{W}$.

We define a sequence $\{\mathbf{V}_i\}_{i \geq 1}$ of pseudovarieties by setting $\mathbf{V}_1 = \mathbf{J}$, and, for $i \geq 1$, $\mathbf{V}_{i+1} = \mathbf{V}_i ** \mathbf{J}$. The main result of [17] is that $\mathbf{DA}$ is the union of the pseudovarieties $\mathbf{V}_i$, and that the variety of languages corresponding to $\mathbf{V}_i$ is the i^{th} level of the alternation hierarchy within $\mathrm{FO}^2[<]$.

2.4 Finite categories

We give a brief account of the tools from the algebraic theory of finite categories needed to prove our main results. The original papers of Tilson [20] and Rhodes and Tilson [12] give a complete and careful exposition of the general theory.

The categories studied in category theory are typically big categories, in which the object class consists of something like all topological spaces, and the arrows are all continuous functions. The work of Tilson [20] showed the utility of studying very small categories in which the object set, as well as each set of arrows between two objects, is finite.

A category $\mathcal{C}$ consists of a set of *objects* $\mathrm{obj}(\mathcal{C})$, a set of *arrows* $\hom(A, B)$ from A to B for all $A, B \in \mathrm{obj}(\mathcal{C})$, and associative partial binary operations $\circ : \hom(A, B) \times \hom(B, C) \to \hom(A, C)$ for all $A, B, C \in \mathrm{obj}(\mathcal{C})$ called *composition*, such that there is an identity in $\hom(A, A)$ for all $A \in \mathrm{obj}(\mathcal{C})$.

In this view, a finite monoid is simply a category with a single object, and a finite category is consequently a generalized finite monoid.

Let A be a finite alphabet, M and N finite monoids with homomorphisms

$$M \xleftarrow{\phi} A^* \xrightarrow{\psi} N,$$

where ϕ maps onto M. We will define a finite category, which we call the *kernel category* $\ker(\psi \circ \phi^{-1})$. [1]

The *objects* of $\ker(\psi \circ \phi^{-1})$ are pairs $(n_1, n_2) \in N \times N$. The *arrows* are represented by triples

$$(n_1, n_2) \xrightarrow{u} (n_1', n_2'),$$

where $u \in A^*$, $n_1' = n_1 \cdot \psi(u)$ and $\psi(u) \cdot n_2' = n_2$. Whenever we have a pair of consecutive arrows

$$(n_1, n_2) \xrightarrow{u} (n_1', n_2'), (n_1', n_2') \xrightarrow{v} (n_1'', n_2''),$$

then we can define the product arrow

$$(n_1, n_2) \xrightarrow{uv} (n_1'', n_2'').$$

If this were all there were to arrows in the kernel category, we would in general have an infinite set of arrows between two objects. However, we identify two coterminal arrows

$$(n_1, n_2) \xrightarrow{u,u'} (n_1', n_2')$$

if for all $v, w \in A^*$ with $\psi(v) = n_1$, $\psi(w) = n_2'$,

$$\phi(vuw) = \phi(vu'w).$$

It is easy to check that this identification is compatible with the product on consecutive arrows, so the true arrows of $\ker(\psi \circ \phi^{-1})$ are equivalence classes modulo this identification. In particular, the finiteness of M and N implies that there are only finitely many distinct arrows.

If $(n_1, n_2) = (n_1', n_2')$, then any pair of arrows from (n_1, n_2) to itself are consecutive, and thus the set of all such arrows at (n_1, n_2) is a finite monoid, which we denote M_{n_1,n_2}. This is a *base monoid.* Base monoids, then, are just built from words u satisfying $n_1 \cdot \psi(u) = n_1$, and $\psi(u) \cdot n_2 = n_2$, and collapsing modulo the equivalence relation identifying arrows.

The following lemma concerning the structure of the base monoids will be quite useful.

▶ **Lemma 3.** *Let A be a finite alphabet: M, N, N' finite monoids, and consider homomorphisms*

$$M \xleftarrow{\phi} A^* \xrightarrow{\psi} N \xrightarrow{\psi'} N',$$

where ϕ maps onto M. Then every base monoid of $\ker(\psi \circ \phi^{-1})$ *divides some base monoid of* $\ker((\psi'\psi) \circ \phi^{-1})$.

Proof. Let $n_1, n_2 \in N$. We denote by M_1 the base monoid at (n_1, n_2) in $\ker(\psi \circ \phi^{-1})$, and by M_2 the base monoid at $(n_1', n_2') = (\psi'(n_1), \psi'(n_2))$ in $\ker((\psi'\psi) \circ \phi^{-1})$. Set

$$U = \{u \in A^* : n_1 \cdot \psi(u) = n_1, n_2 = \psi(u) \cdot n_2\},$$

$$U' = \{u \in A^* : n_1' \cdot \psi'\psi(u) = n_1', n_2' = \psi'\psi(u) \cdot n_2'\}.$$

[1] The odd notation for the kernel category is used to maintain consistency with the traditional setting for these finite categories. $\psi \circ \phi^{-1}$ is a *relational morphism* from M to N, and Tilson defines these categories for arbitrary relational morphisms, not just those derived from morphisms of the free monoid.

U and U' are submonoids of A^*, and $U \subseteq U'$. M_1 and M_2 are the quotients of U and U' by the congruences identifying equivalent arrows in the respective categories. Let $u, u' \in U$ represent equivalent arrows of M_2, and suppose $v, w \in A^*$ are such that $\psi(v) = n_1$, $\psi(w) = n_2$. Then $\psi'\psi(v) = n_1'$, $\psi'\psi(w) = n_2'$, so by equivalence in M_2 we have $\phi(vuw) = \phi(vu'w)$. But this means that u and u' represent equivalent arrows in M_1, so M_1 is a quotient of the image of U in M_2. Thus $M_1 \prec M_2$. ◀

It is worth keeping in mind the somewhat counterintuitive message of this lemma: The category $\ker(\psi \circ \phi^{-1})$ is *bigger* (it has more objects) than $\ker((\psi'\psi) \circ \phi^{-1})$ but its base monoids are *smaller*.

The reason for the construction of the kernel category in [12] is its relation to two-sided semidirect products. Roughly speaking, $M \in \mathbf{V} ** \mathbf{W}$ if and only if there exists $\psi : A^* \to N \in \mathbf{W}$ such that the category $\ker(\psi \circ \phi^{-1})$ is 'globally in $\mathbf{V}$'. We will not define this precisely, but instead note without proof one consequence, namely that if $M \in \mathbf{V} ** \mathbf{W}$, then $\ker(\psi \circ \phi^{-1})$ satisfies a weaker condition of being 'locally in $\mathbf{V}$':

▶ **Proposition 4.** *Let $\phi : A^* \to M$ be a homomorphism mapping onto M. If $M \in \mathbf{V} ** \mathbf{W}$, then there is a homomorphism $\psi : A^* \to N \in \mathbf{W}$ such that each base monoid of $\ker(\psi \circ \phi^{-1})$ is in $\mathbf{V}$.*

3 A local-global theorem for categories

In general, the converse of Proposition 4 is false. This section is devoted to establishing an important instance in which it is true, namely when $\mathbf{W} = \mathbf{J}$.

▶ **Theorem 5.** *Let A be a finite alphabet, M and N finite monoids with $N \in \mathbf{J}$ and homomorphisms*

$$M \xleftarrow{\phi} A^* \xrightarrow{\psi} N,$$

where ϕ maps onto M. Suppose $\mathbf{V}$ is a pseudovariety of finite monoids with $\mathbf{J_1} \subseteq \mathbf{V}$. If every base monoid of $\ker(\psi \circ \phi^{-1})$ is in $\mathbf{V}$, then $M \in \mathbf{V} ** \mathbf{J}$.

Proof. It follows from Theorem 2 that for some $k > 0$, ψ factors through the homomorphism $A^* \to A^*/\sim_k$ identifying two words that have the same subwords up to length k. By Lemma 3 we may assume that ψ *is* this homomorphism, and that $N = A^*/\sim_k$. In particular, if $w \in A^*$, then we can represent $\psi(w)$ as the set of subwords of w of length no more than k.

The set $\mathcal{P}(N \times N)$ of subsets of $N \times N$ forms an idempotent and commutative monoid with union as the operation, and hence belongs to $\mathbf{J_1} \subseteq \mathbf{V}$. Let $\Sigma = N \times A \times N$ and let $h_U : \Sigma^* \to \mathcal{P}(N \times N)$ be the homomorphism defined by mapping $\sigma = (P, a, S)$ to $\{(P\psi(a), S)\}$ for each $\sigma \in \Sigma$. Given $P, S \in N$, define a homomorphism $h_{P,S} : \Sigma^* \to M_{P,S}$ by mapping $(P', a, S') \in \Sigma$ to the arrow class of $(P, S) \xrightarrow{a} (P, S)$ if $P = P' = P\psi(a), \psi(a)S = S' = S$, and to $1 \in M_{P,S}$ otherwise. Finally, set M' to be the direct product

$$M' = \mathcal{P}(N \times N) \times \prod_{(P,S) \in N \times N} M_{P,S},$$

and set

$$h = h_U \times \prod_{(P,S) \in N \times N} h_{P,S}.$$

By our hypothesis $M' \in \mathbf{V}$.

Let $w, w' \in A^*$, with $\psi(w) = \psi(w')$ and $h(\tau_\psi(w)) = h(\tau_\psi(w'))$. We will show $\phi(w) = \phi(w')$, which gives the result.

We will look at the paths through $\ker(\psi \circ \phi^{-1})$ traced out by w and w'. Since $\psi(w) = \psi(w')$, the two paths are coterminal, beginning at the object $(1, \psi(w))$ and ending at $(\psi(w), 1)$. Let the i^{th} letter of w be a_i, then the i^{th} arrow on this path is the class of

$$(P_{i-1}, S_{i-1}) \xrightarrow{a_i} (P_i, S_i),$$

where P_j is $\psi(u)$ for the prefix $u = a_1 \cdots a_j$ of length j of w, and likewise $S_j = \psi(v)$ for the suffix $v = a_{j+1} \cdots a_{|w|}$. Let

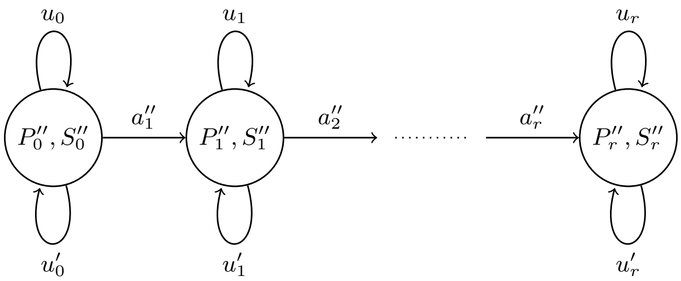
$$(P, S) \xrightarrow{a_i} (P', S')$$

be on the path traced by w, then we have $P \subseteq P'$ and $S' \subseteq S$. Either $P = P'$ and $S = S'$, in which case this arrow belongs to one of the base monoids, or at least one of the inclusions is proper. Since $h_U(\tau_\psi(w)) = h_U(\tau_\psi(w'))$ and $\psi(w) = \psi(w')$, the same pairs $(P, S), (P', S')$ must occur in the path traced by w'. Because of the inclusions, they must occur in the same relative order in this path, with (P, S) preceding (P', S'). They also must be adjacent in this path, since if there were a third pair (P'', S'') between them, we would have

$$P \subseteq P'' \subseteq P', S' \subseteq S'' \subseteq S,$$

so this new pair would have to occur in the original path traced by w, strictly between (P, S) and (P', S'). Finally, the letter a labeling the arrow joining these two objects in the respective paths is completely determined by (P, S) and (P', S'). This is because at least one of the two inclusions $P \subseteq P'$ and $S' \subseteq S$ is proper. Assume without loss of generality that the first of these is a proper inclusion. Then P' contains a word that is not in P, and the last letter of this word is a.

Thus our two paths are depicted by the diagram below:



The paths traverse exactly the same sequence of distinct objects

$$(1, \psi(w)) = (P''_0, S''_0), (P''_1, S''_1), \ldots, (P''_r, S''_r) = (\psi(w), 1).$$

The arrow joining (P''_{j-1}, S''_{j-1}) and (P''_j, S''_j) in both these paths is the same letter a''_j. For $j = 0, \ldots, r$ each path contains a loop at (P''_j, S''_j) labeled by a factor u_j of w in one path, and a factor u'_j in the other path. We have

$$w = u_0 a''_1 u_1 \cdots a''_r u_r,$$

$$w' = u'_0 a''_1 u'_1 \cdots a''_r u'_r.$$

Let $w_0 = w$, $w_{r+1} = w'$ and for $j = 1, \ldots, r$, let

$$w_j = u'_0 a''_1 \cdots u'_{j-1} a''_j u_j \cdots a''_r u_r.$$

In other words, we transform w into w' one step at a time, changing each u_j in succession to u'_j. We claim that at each step, $\phi(w_j) = \phi(w_{j+1})$, so that we will get $\phi(w) = \phi(w')$, as required. Let $(P, S) = (P''_j, S''_j)$. By the definition of $h_{P,S}$, the image in the base monoid $M_{P,S}$ of a word under $h_{P,S} \circ \tau_\psi$ involves only the letters where the prefix under ψ maps to P and suffix maps to S; in our case these are the letters of u_j and u'_j, respectively. So since $h(\tau_\psi(w)) = h(\tau_\psi(w'))$, $(P, S) \xrightarrow{u_j, u'_j} (P, S)$ are equivalent arrows. Thus

$$\phi(w_j) = \phi(u'_0 a''_1 \cdots u'_{j-1} a''_j u_j a''_{j+1} \cdots u_r) = \phi(u'_0 a''_1 \cdots u'_{j-1} a''_j u'_j a''_{j+1} \cdots u_r) = \phi(w_{j+1}).$$

◀

As we mentioned in the introduction, our argument is essentially the one used to prove a more general result, due to Tilson [20], that every category divides a direct product of its strongly connected components. In the special case that we consider, these components reduce to single objects and hence are the base monoids of the category. The hypothesis $\mathbf{J_1} \subseteq \mathbf{V}$ in the statement of Theorem 5 is actually not necessary, so long as $\mathbf{V}$ is nontrivial: If $\mathbf{V}$ is a pseudovariety of monoids that does not contain $\mathbf{J_1}$, then every member of $\mathbf{V}$ is a group, and it is known that the converse of Proposition 4 holds when $\mathbf{V}$ contains only nontrivial groups; this also follows from results in [20]. We do not require this fact in our main application to the alternation hierarchy.

4 Effective characterization of levels of the alternation hierarchy

We now define a sequence of identities that will allow us to characterize the varieties $\mathbf{V}_n$. We set

$$u_1 = (x_1x_2)^\omega, v_1 = (x_2x_1)^\omega,$$

and for $n \geq 1$,

$$u_{n+1} = (x_1 \cdots x_{2n}x_{2n+1})^\omega u_n (x_{2n+2}x_1 \cdots x_{2n})^\omega,$$

$$v_{n+1} = (x_1 \cdots x_{2n}x_{2n+1})^\omega v_n (x_{2n+2}x_1 \cdots x_{2n})^\omega.$$

▶ **Theorem 6.** *Let $n \geq 1$. $M \in \mathbf{V}_n$ if and only if $M \models (u_n = v_n)$, and $M \models (x^\omega = xx^\omega)$.*

As we remarked above, when a pseudovariety $\mathbf{V}$ is defined by a finite set of identities of this type, one can decide membership in $\mathbf{V}$. Since the levels of the alternation hierarchy in $\mathrm{FO}^2[<]$ are the varieties of languages corresponding to the $\mathbf{V}_i$, the alternation depth of a language in $\mathrm{FO}^2[<]$ is effectively computable.

Proof. The 'only if' part (the identities hold in $\mathbf{V}_n$) is proved in [17], so we will just give the proof of the 'if' part (sufficiency of the identities).

We prove the theorem by induction on n. It is well known that the identities $u_1 = v_1, x^\omega = xx^\omega$ characterize $\mathbf{V}_1 = \mathbf{J}$.

So we assume $n > 1$ and suppose that M is an aperiodic monoid such that $M \models (u_n = v_n)$. We let ϕ, ψ be as in the previous section, so that ϕ is any morphism mapping onto $M \in \mathbf{V}_n$, and ψ depends on the choice of a subword length K. We will show that if K is chosen to be a large enough value, then each base monoid $M_{P,S}$ of the category $\ker(\psi \circ \phi^{-1})$ satisfies the identity $u_{n-1} = v_{n-1}$. By the inductive hypothesis and since $M_{P,S}$ is aperiodic, this implies that each $M_{P,S}$ belongs to $\mathbf{V}_{n-1}$, and thus by Theorem 5, $M \in \mathbf{V}_{n-1} ** \mathbf{J} = \mathbf{V}_n$.

We let $x_1, \ldots, x_{2(n-1)}$ be any elements of $M_{P,S}$. Thus, each x_i is represented by a triple $(P, S) \xrightarrow{w_i} (P, S)$, where $w_i \in A^*$, $P \cdot \psi(w_i) = P$, $\psi(w_i) \cdot S = S$.

We construct words $W_{n-1}, W'_{n-1} \in A^*$ by replacing each x_i in u_{n-1} (respectively v_{n-1}) by w_i. We will think of ω in these strings as representing a finite exponent N such that $x^N = x^{N+1}$ for all $x \in M$, and hence for all $x \in M_{P,S}$, since each base monoid $M_{P,S}$ divides M. Thus if $n > 2$,

$$W_{n-1} = (w_1 w_2 \cdots w_{2n-3})^N W_{n-2} (w_{2n-2} w_1 \cdots w_{2n-4})^N,$$

$$W'_{n-1} = (w_1 w_2 \cdots w_{2n-3})^N W'_{n-2} (w_{2n-2} w_1 \cdots w_{2n-4})^N.$$

In the special case $n = 2$, we have $W_1 = (w_1 w_2)^N$, $W'_1 = (w_2 w_1)^N$.

If $w \in A^*$, we denote by $\alpha(w)$ the set of letters occurring in A^*. We also denote by B the set $\alpha(W_{n-1}) = \alpha(W'_{n-1})$. Let $z, y \in A^*$ with $\psi(z) = P$, $\psi(y) = S$.

▶ **Lemma 7.** *If $K > |M| \cdot (|A|^2 + |A|)/2$, then z has a suffix z' with a factorization $z' = z_1 z_2 \cdots z_{|M|}$ where*

$$B \subseteq \alpha(z_1) = \alpha(z_2) = \cdots = \alpha(z_{|M|}),$$

and, likewise, y has a prefix y' with a factorization $y' = y_1 y_2 \cdots y_{|M|}$ where

$$B \subseteq \alpha(y_1) = \alpha(y_2) = \cdots = \alpha(y_{|M|}).$$

Assuming the lemma, we will now complete the proof of Theorem 6. Since $M \models (u_n = v_n)$, we obtain $M \models (xy)^\omega (yx)^\omega (xy)^\omega = (xy)^\omega$ by setting $x_1 = x$, $x_2 = y$, and $x_k = 1$ for $k > 2$. Thus $M \in \mathbf{DA}$.

We can write $z = z'' z'$, where $z' = z_1 \ldots z_{|M|}$ has a factorization as in Lemma 7. By the standard pumping argument, it follows that there are indices $i \leq j$ such that $\phi(z_1 \ldots z_{i-1}) \phi(z_i \cdots z_j) = \phi(z_1 \ldots z_{i-1})$, and thus

$$\phi(z_1 \ldots z_{i-1}) \phi(z_i \cdots z_j)^\omega = \phi(z_1 \ldots z_{i-1}).$$

If we now set

$$e = \phi(z_i \cdots z_j)^\omega$$

$$s_{2n-1} = e \cdot \phi(z_{j+1} \cdots z_{|M|}), \text{ and } s_i = \phi(w_i) \text{ for } i < 2n-1,$$

we obtain, from the identity $e \cdot M_e \cdot e = e$,

$$\begin{aligned}
\phi(z') &= \phi(z_1 \ldots z_{i-1}) \cdot \phi(z_{j+1} \cdots z_{|M|}) \\
&= \phi(z_1 \ldots z_{i-1}) e \cdot e \phi(z_{j+1} \cdots z_{|M|}) \\
&= \phi(z_1 \ldots z_{i-1}) e \cdot \left(\phi(z_{j+1} \cdots z_{|M|}) (s_1 \cdots s_{2n-1})^{\omega - 1} s_1 \cdots s_{2n-2} \right) \cdot e \phi(z_{j+1} \cdots z_{|M|}) \\
&= \phi(z_1 \ldots z_{i-1}) e \cdot \phi(z_{j+1} \cdots z_{|M|}) (s_1 \cdots s_{2n-1})^\omega \\
&= \phi(z') \cdot (s_1 \cdots s_{2n-1})^\omega.
\end{aligned}$$

The third equality above holds because by Lemma 7, $z_i \cdots z_j$ contains all the letters that occur in the z_k and the w_k, and hence all the values we inserted between occurrences of e belong to M_e.

Similarly, using the part of Lemma 7 concerning the prefix of y, we find a value s_{2n} such that $\phi(y') = (s_{2n} s_1 \cdots s_{2n-2})^\omega \phi(y')$. Since $M \models (u_n = v_n)$ we obtain

$$\begin{aligned}
\phi(z W_{n-1} y) &= \phi(z'') \phi(z') \phi(W_{n-1}) \phi(y') \phi(y'') \\
&= \phi(z'') \phi(z') (s_1 \cdots s_{2n-1})^\omega \phi(W_{n-1}) (s_{2n} s_1 \cdots s_{2n-2})^\omega \phi(y') \phi(y'') \\
&= \phi(z'') \phi(z') (s_1 \cdots s_{2n-1})^\omega \phi(W'_{n-1}) (s_{2n} s_1 \cdots s_{2n-2})^\omega \phi(y') \phi(y'') \\
&= \phi(z'') \phi(z') \phi(W_{n-1}) \phi(y') \phi(y'') \\
&= \phi(z W'_{n-1} y).
\end{aligned}$$

But this means that $M_{P,S} \models (u_{n-1} = v_{n-1})$, as we required. ◀

We now turn to the proof of Lemma 7.

Proof of Lemma 7. By symmetry, we only need to treat the part concerning the suffix of z. Whenever we need to emphasize the dependence of ψ on the chosen subword length m, we will write it as ψ_m. Recall that $\psi(z) = \psi_K(z)$ is the set of subwords of length no more than K in z, and that $\psi(zb) = \psi(z)$ for all $b \in B$.

We will show that if $B \subseteq \alpha(z)$ and $\psi_T(zb) = \psi_T(z)$ for all $b \in B$, where

$$T = |M| \cdot (k^2 + k)/2,$$

and $k = |\alpha(z)|$, then z contains a suffix with the required properties. This gives the lemma, because $\psi_K(zb) = \psi_K(z)$ implies $\psi_T(zb) = \psi_T(z)$ for any $\alpha(z) \subseteq A$.

The proof is by induction on $|\alpha(z)|$. The base case occurs when $\alpha(z) = B$. Let $B = \{b_1, \ldots b_r\}$. By repeated application of $\psi_T(zb_i) = \psi_T(z)$ we find $(b_1 \cdots b_r)^{|M|}$, which has length $|M||B| \leq |M|(|B|^2 + |B|)/2$, is a subword of z. In this case we can simply take $z' = z$ and factor $z = z_1 \cdots z_{|M|}$, where each z_i contains one of the factors $b_1 \cdots b_r$ as a subword.

We thus suppose that $\alpha(z) = A'$ contains B as a proper subset. Let $N = |A'|$. We look at a subword of maximal length $t_1 \cdots t_p$ of z such that $\alpha(t_i) = A'$. We must have $p \geq 1$. If $p \geq |M|$, we can again take $z' = z$ and factor z as $z_1 \cdots z_{|M|}$, where each z_i contains t_i as a subword. If $p < |M|$, we let $s = |A'|$, then we write

$$t_1 \cdots t_p = a_1 a_2 \cdots a_{ps}, \text{ and } z = z_0 a_1 z_1 \cdots a_{ps} z_{ps}.$$

We further suppose that this factorization represents the leftmost occurrence of $a_1 \cdots a_{ps}$ as a subword of z, in other words that z_{ps} has maximum possible length for this property. Note that $\alpha(z_{ps})$ is a strict subset of A', for otherwise z would have contained a longer subword $t_1 \cdots t_{p+1}$ with $\alpha(t_i) = A'$. Thus $|\alpha(z_{ps})| \leq N - 1$. Set $T = |M| \cdot ((N-1)^2 + (N-1))/2$. We must have $\psi_T(z_{ps}b) = \psi_T(z_{ps})$ for all $b \in B$. If not, there is a subword u of z_{ps} of length less than T such that ub is not a subword of z_{ps}. However $t_1 \cdots t_p ub$ has length no more than

$$(|M| - 1) \cdot N + T < |M| \cdot (N + ((N-1)^2 + (N-1))/2) = |M| \cdot (N^2 + N)/2,$$

and is accordingly a subword of z, and thus there is a strictly earlier occurrence of $t_1 \cdots t_p$ as a subword of z, a contradiction. We can thus apply the inductive hypothesis to z_{ps} and conclude that z_{ps} contains a suffix of the required type. ◀

5 Collapse of the hierarchy

In the original model-theoretic study of the alternation hierarchy in $\mathrm{FO}^2[<]$, Weis and Immerman [21] and also Kufleitner and Weil [9] showed that while the hierarchy is strict, it collapses for each fixed-size alphabet. An algebraic proof of strictness was given in [17], using the identities that form the subject of the present paper. We can use similar techniques to prove the collapse result.

▶ **Theorem 8.** *Let $n > 0$. If $M \in \mathbf{DA}$ is generated by n elements, then $M \in \mathbf{V}_n$.*

The proof, which we omit, uses our main result Theorem 6 to conclude that if $M \in \mathbf{DA}$, then $M \models (u_N = v_N)$ for some N. We then use identity manipulation in $\mathbf{DA}$ to show that this implies $M \models (u_n = v_n)$, where n is the number of generators.

In particular for any fixed alphabet the quantifier alternation hierarchy collapses.

▶ **Corollary 9.** *Any language over a k-letter alphabet definable by a two-variable sentence is definable by one in which the number of quantifier blocks is k.*

6 General decidability results

Here we show that for arbitrary pseudovarieties $\mathbf{V}$, the operation $\mathbf{V} \mapsto \mathbf{V} ** \mathbf{J}$ preserves decidability. This of course implies our result (a consequence of Theorem 6) that the varieties $\mathbf{V}_j$ are all decidable, but Theorem 6 is a sharper result, since it gives explicit identities. As we remarked in the introduction, the general decidability result was originally proved by Steinberg [16], but not previously published. Our approach has the advantages both of being relatively elementary, and yielding explicit bounds on the complexity of membership testing.

We suppose that $\phi : A^* \to M$ is a surjective homomorphism onto a finite monoid. Let $N > 0$, we denote by $\ker_N \phi$ the category $\ker(\psi_N \circ \phi^{-1})$, where ψ_N is the natural projection of A^* onto the quotient $A^*/\sim_N$. We set

$$K = |M| \cdot (|A|^2 + |A|)/2$$

as in the statement of Lemma 7. With these notations we have:

▶ **Theorem 10.** *Let $\mathbf{V}$ be a pseudovariety of monoids. $M \in \mathbf{V} ** \mathbf{J}$ if and only if every base monoid of $\ker_K \phi$ is in $\mathbf{V}$.*

We omit the proof. The idea is this: By Proposition 4, if $M \in \mathbf{V} ** \mathbf{J}$, then there is some K' such that all the base monoids of $\ker_{K'} \phi$ are in $\mathbf{V}$. We use Lemma 7 to show that if $K' > K$, then every base monoid of $\ker_K \phi$ divides a direct product of base monoids in $\ker_{K'} \phi$, so the result follows from Theorem 5.

We can effectively compute all the objects and arrow classes of $\ker_K \phi$ from ϕ, and we can also take $A = M$ and ϕ to be the extension of the identity map on M to A^*. The theorem thus immediately implies:

▶ **Corollary 11.** *If $\mathbf{V}$ is a decidable pseudovariety of finite monoids, then so is $\mathbf{V} ** \mathbf{J}$.*

7 Conclusion

We have shown that the identities given in [17] indeed characterize $\mathbf{V}_n$. There is, of course, a one-sided semidirect product, which has been much more thoroughly studied. Our results, and their proofs, can all be adapted to one-sided products, with little modification. In this case, the hierarchy collapses at the second level: $\mathbf{J} * \mathbf{J} * \mathbf{J} = \mathbf{J} * \mathbf{J}$. (This fact is not new. It has long been known that the closure of $\mathbf{J}$ under one-sided products is the pseudovariety $\mathbf{R}$ of $\mathcal{R}$-trivial monoids, and Brzozowski and Fich [3] showed $\mathbf{R} = \mathbf{J}_1 * \mathbf{J}$.)

In their paper, Kufleitner and Weil [9] give a completely different characterization of the levels of $\mathrm{FO}^2[<]$. It would be nice to see a direct connection between these two approaches.

Acknowledgements We are grateful to Manfred Kufleitner, Benjamin Steinberg, and Pascal Weil for detailed discussions of this work, and to the anonymous referees for their helpful suggestions.

References

1 Jorge Almeida. *Finite Semigroups and Universal Algebra.* Series in Algebra. World Scientific, 1994.

2 Jorge Almeida and Pascal Weil. Profinite categories and semidirect products. *Journal of Pure and Applied Algebra*, 123(1-3):1–50, 1998.

3 Janusz A. Brzozowski and Faith E. Fich. Languages of R-trivial monoids. *J. Comput. Syst. Sci.*, 20(1):32–49, 1980.

4 Samuel Eilenberg. *Automata, Languages, and Machines Vol. 2.* Pure and applied mathematics. Academic Press, 1976.

5 Neil Immerman and Dexter Kozen. Definability with bounded number of bound variables. *Inf. Comput.*, 83(2):121–139, 1989.

6 Johan A. W. Kamp. *Tense logic and the theory of linear order.* PhD thesis, University of California, Los Angeles, 1968.

7 Kenneth Krohn, Richard Mateosian, and John Rhodes. Methods of the algebraic theory of machines. I: Decomposition theorem for generalized machines; properties preserved under series and parallel compositions of machines. *J. Comput. Syst. Sci.*, 1(1):55–85, 1967.

8 Manfred Kufleitner and Pascal Weil. On FO^2 quantifier alternation over words. In *Mathematical Foundations of Computer Science*, pages 513–524, 2009.

9 Manfred Kufleitner and Pascal Weil. The FO^2 alternation hierarchy is decidable. In *Computer Science Logic*, pages 426–439, 2012.

10 R. McNaughton and S. Papert. *Counter-free automata.* M.I.T. Press research monographs. M.I.T. Press, 1971.

11 Jean-Éric Pin. *Varieties of formal languages.* North Oxford Academic, 1986.

12 John L. Rhodes and Bret Tilson. The kernel of monoid morphisms. *Journal of Pure and Applied Algebra*, 62:227–268, 1989.

13 Marcel Paul Schützenberger. A remark on finite transducers. *Information and Control*, 4(2-3):185–196, 1961.

14 Marcel Paul Schützenberger. Sur le produit de concatenation non ambigu. *Semigroup Forum*, 13:47–75, 1976. 10.1007/BF02194921.

15 Imre Simon. Piecewise testable events. In *Automata Theory and Formal Languages*, pages 214–222, 1975.

16 Benjamin Steinberg. *Decidability and Hyperdecidability of Joins of Pseudovarieties.* PhD thesis, University of California at Berkeley, 1998.

17 Howard Straubing. Algebraic characterization of the alternation hierarchy in $\mathrm{FO}^2[<]$ on finite words. In *Computer Science Logic*, pages 525–537, 2011.

18 Pascal Tesson and Denis Therien. Diamonds are forever: The variety DA. In *Semigroups, Algorithms, Automata and Languages, Coimbra (Portugal) 2001*, pages 475–500. World Scientific, 2002.

19 Denis Thérien and Thomas Wilke. Over words, two variables are as powerful as one quantifier alternation. In *STOC*, pages 234–240, 1998.

20 Bret Tilson. Categories as algebra: An essential ingredient in the theory of monoids. *Journal of Pure and Applied Algebra*, 48(1-2):83–198, 1987.

21 Philipp Weis and Neil Immerman. Structure theorem and strict alternation hierarchy for FO^2 on words. *Logical Methods in Computer Science*, 5(3), 2009.