Block Products and Nesting Negations in FO²

Lukas Fleischer[1], Manfred Kufleitner[1,2], and Alexander Lauser[1]

[1] Formale Methoden der Informatik, Universität Stuttgart, Germany[⋆]
[2] Fakultät für Informatik, Technische Universität München, Germany
fleiscls@studi.informatik.uni-stuttgart.de
kufleitn@in.tum.de
lauser@fmi.uni-stuttgart.de

Abstract. The alternation hierarchy in two-variable first-order logic $FO^2[<]$ over words was recently shown to be decidable by Kufleitner and Weil, and independently by Krebs and Straubing. In this paper we consider a similar hierarchy, reminiscent of the half levels of the dot-depth hierarchy or the Straubing-Thérien hierarchy. The fragment Σ_m^2 of FO^2 is defined by disallowing universal quantifiers and having at most $m - 1$ nested negations. One can view Σ_m^2 as the formulas in FO^2 which have at most m blocks of quantifiers on every path of their parse tree, and the first block is existential. Thus, the mth level of the FO^2-alternation hierarchy is the Boolean closure of Σ_m^2. We give an effective characterization of Σ_m^2, *i.e.*, for every integer m one can decide whether a given regular language is definable by a two-variable first-order formula with negation nesting depth at most m. More precisely, for every m we give ω-terms U_m and V_m such that an FO^2-definable language is in Σ_m^2 if and only if its ordered syntactic monoid satisfies the identity $U_m \leqslant V_m$. Among other techniques, the proof relies on an extension of block products to ordered monoids.

1 Introduction

The study of logical fragments over words has a long tradition in computer science. The seminal Büchi-Elgot-Trakhtenbrot Theorem from the early 1960s states that a language is regular if and only if it is definable in monadic second-order logic [1,5,32]. A decade later, in 1971, McNaughton and Papert showed that a language is definable in first-order logic if and only if it is star-free [17]. Combining this result with Schützenberger's famous characterization of star-free languages in terms of finite aperiodic monoids [21] shows that it is decidable whether a given regular language is first-order definable. Since then, many logical fragments have been investigated, see *e.g.* [3,25] for overviews.

The motivation for such results is two-fold. First, restricted fragments often yield more efficient algorithms for computational problems such as satisfiability or separability. Second, logical fragments give rise to a descriptive complexity:

The simpler the fragment to define a language, the simpler the language. This approach can help in understanding the rich structure of regular languages.

Logical fragments are usually defined by restricting some resources in formulas. The three most natural restrictions are the quantifier depth (*i.e.*, the number of nested quantifiers), the alternation depth (*i.e.*, the number alternations between existential and universal quantification), and the number of variables. With respect to decidability questions regarding definability, quantifier depth is not very interesting since for fixed quantifier depth only finitely many languages are definable (which immediately yields decidability), see *e.g.* [4]. The situation with alternation in first-order logic is totally different: Only the very first level (*i.e.*, no alternation) is known to be decidable [8,23]. By a result of Thomas [31] the alternation hierarchy in first-order logic is tightly connected with the dot-depth hierarchy [2] or the Straubing-Thérien hierarchy [24,29], depending on the presence or absence of the successor predicate. Some progress in the study of the dot-depth hierarchy and the Straubing-Thérien hierarchy was achieved by considering the half-levels. For example, the levels ½ and ³⁄₂ in each of the two hierarchies are decidable [6,18,19]. The half levels also have a counterpart in the alternation hierarchy of first-order logic by requiring existential quantifiers in the first block. Another point of view of the same hierarchy is to disallow universal quantifiers and to restrict the number of nested negations.

Regarding the number of variables, Kamp showed that linear temporal logic is expressively complete for first-order logic over words [7]. Since every modality in linear temporal logic can be defined using three variables, first-order logic with only three different names for the variables (denoted by FO^3) defines the same languages as full first-order logic. This result is often stated as $FO^3 = FO$. Allowing only two variable names yields the proper fragment FO^2 of first-order logic. Thérien and Wilke [30] showed that a language is FO^2 definable if and only if its syntactic monoid belongs to the variety **DA** and, since the latter is decidable, one can effectively check whether a given regular language is FO^2-definable. For further information on the numerous characterizations of FO^2 we refer to [3,28].

Inside FO^2, the alternation depth is also a natural restriction. One difference to full first-order logic is that one cannot rely on prenex normal forms as a simple way of defining the alternation depth. Weil and the second author gave an effective algebraic characterization of the mth level FO^2_m of this hierarchy. More precisely, they showed that it is possible to ascend the FO^2-alternation hierarchy using so-called Mal'cev products [15] which in this particular case preserve decidability. There are two main ingredients in the proof. The first one is a combinatorial tool known as *rankers* [33] or *turtle programs* [22], and the second is a relativization property of two-variable first-order logic. These two ingredients are then combined using a proof method introduced in [10]. Krebs and Straubing gave another decidable characterization of FO^2_m in terms of identities of ω-terms using completely different techniques [9,26]; their proof relies on so-called block products.

In this paper we consider the half-levels Σ^2_m of the FO^2-alternation hierarchy. A language is definable in Σ^2_m if and only if it is definable in FO^2 without

universal quantifiers and with at most $m-1$ nested negations. It is easy to see that one can avoid negations of atomic predicates. One can think of Σ_m^2 as those FO^2-formulas which on every path of their parse tree have at most m quantifier blocks, and the outermost block is existential. The main contribution of this paper are ω-terms U_m and V_m such that an FO^2-definable language is Σ_m^2-definable if and only if its ordered syntactic monoid satisfies $U_m \leqslant V_m$. For a given regular language it is therefore decidable whether it is definable in Σ_m^2 by first checking whether it is FO^2-definable and if so, then verifying whether $U_m \leqslant V_m$ holds in its ordered syntactic monoid. Moreover, for every FO^2-definable language L one can compute the smallest integer m such that L is definable in Σ_m^2.

The proof step from the identities to logic is a refinement of the approach of Weil and the second author [15] which in turn uses a technique from [10, Section IV]. While the proof method in [10] is quite general and can be applied for solving various other problems [11,12,13,14], it relies on closure under negation. A very specific modification is necessary in order to get the scheme working in the current situation.

The proof for showing that Σ_m^2 satisfies the identity $U_m \leqslant V_m$ is an adaptation of Straubing's proof [26] to ordered monoids. Straubing's proof relies on two-sided semidirect products and the block product principle. We partially extend both tools to ordered monoids. To the best of our knowledge, this extension does not yet appear in the literature. The attribute *partially* is due to the fact that only the first factor in two-sided semidirect products (as used in this paper) is ordered while the second factor is an unordered monoid. As shown by Pin and Weil in the case of one-sided semidirect products [20], one could use ordered alphabets for further extending this approach. We refrain from this in order to focus on the presentation of our main result.

2 Preliminaries

The *free monoid* A^* is the set of finite words over A equipped with concatenation and the empty word ε as neutral element. Let $u = a_1 \cdots a_k$ with $a_i \in A$ be a finite word. The *alphabet* (also known as the *content*) of u is $\mathrm{alph}(u) = \{a_1, \ldots, a_k\}$, its *length* is $|u| = k$, and the *positions* of u are $1, \ldots, k$. We say that i is an *a-position* of u if $a_i = a$. The word u is a *(scattered) subword* of w if $w \in A^* a_1 \cdots A^* a_k A^*$.

First-order logic. We consider first-order logic $\mathrm{FO} = \mathrm{FO}[<]$ over finite words. The syntax of FO-formulas is

$$\varphi ::= \top \mid \bot \mid \lambda(x) = a \mid x = y \mid x < y \mid \neg\varphi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \exists x\, \varphi$$

where $a \in A$ is a letter, and x and y are variables. We consider universal quantifiers $\forall x\, \varphi$ as an abbreviation of $\neg \exists x\, \neg\varphi$, and $x \leqslant y$ is a shortcut for $(x = y) \vee (x < y)$. The atomic formulas $\top$ and $\bot$ are *true* and *false*, respectively. Variables are interpreted as positions of a word, and $\lambda(x) = a$ is true if x is an a-position. The semantics of the other constructs is as usual; in particular, $\exists x\, \varphi$ means that there exists a position x which makes φ true, and $x < y$ means that position x is (strictly) smaller than position y. We write $\varphi(x_1, \ldots, x_\ell)$ for a formula φ if at most

the variables x_i appear freely in φ; and we write $u, p_1, \ldots, p_\ell \models \varphi(x_1, \ldots, x_\ell)$ if φ is true over u when x_i is interpreted as p_i. A *sentence* is a formula without free variables. A first-order sentence φ *defines* the language $L(\varphi) = \{u \in A^* \mid u \models \varphi\}$, and a language is *definable* in a first-order fragment $\mathcal{F}$ if it is defined by some sentence in $\mathcal{F}$.

The formulas φ_m in the mth level Σ_m of the *negation nesting* hierarchy in FO are defined as follows:

$$\varphi_m \ ::= \ \varphi_{m-1} \ \mid \ \neg\varphi_{m-1} \ \mid \ \varphi_m \vee \varphi_m \ \mid \ \varphi_m \wedge \varphi_m \ \mid \ \exists x \, \varphi_m$$
$$\varphi_0 \ ::= \ \top \ \mid \ \bot \ \mid \ \lambda(x) = a \ \mid \ x = y \ \mid \ x < y \ \mid \ \neg\varphi_0 \ \mid \ \varphi_0 \vee \varphi_0 \ \mid \ \varphi_0 \wedge \varphi_0$$

This means, for $m \geqslant 1$ the formulas in Σ_m have at most $m-1$ nested negations over quantifier-free formulas φ_0. Using De Morgan's laws and the following equivalences, one can avoid negations in quantifier-free formulas for fixed alphabet A:

$$\lambda(x) \neq a \ \equiv \ \bigvee_{b \in A \setminus \{a\}} \lambda(x) = b$$
$$x \neq y \ \equiv \ (x < y) \vee (y < x)$$
$$\neg(x < y) \ \equiv \ (x = y) \vee (y < x)$$

Also note that, up to logical equivalence, our definition of Σ_m coincides with the more common definition in terms of formulas in prenex normal form with at most m blocks of quantifiers which start with an existential block. This can be seen by the usual procedure of renaming the variables and successively moving quantifiers outwards.

The two-variable fragment FO^2 of first-order logic uses (and reuses) only two different variables, say x and y. Combining FO^2 and Σ_m yields the fragment Σ_m^2. That is, we have $\varphi \in \Sigma_m^2$ if both $\varphi \in \Sigma_m$ and $\varphi \in \text{FO}^2$. This also justifies the notation Σ_m^2 which inherits the symbol as well as the subscript from Σ_m and the exponent from FO^2. The Boolean closure of Σ_m^2 is the mth level FO_m^2 of the *alternation hierarchy* within FO^2.

Ordered monoids. *Green's relations* are an important tool in the study of finite monoids. For $x, y \in M$ let $x \leqslant_{\mathcal{R}} y$ if $xM \subseteq yM$, and let $x \leqslant_{\mathcal{L}} y$ if $Mx \subseteq My$. We write $x \, \mathcal{R} \, y$ if both $x \leqslant_{\mathcal{R}} y$ and $y \leqslant_{\mathcal{R}} x$; and we set $x <_{\mathcal{R}} y$ if $x \leqslant_{\mathcal{R}} y$ but not $x \, \mathcal{R} \, y$. The relations $\mathcal{L}$ and $<_{\mathcal{L}}$ are defined similarly. An element $x \in M$ is *idempotent* if $x^2 = x$. For every finite monoid M there exists an integer $\omega_M \geqslant 1$ such that x^{ω_M} is the unique idempotent power generated by $x \in M$. If the reference to M is clear from the context, we simply write ω instead of ω_M.

An *ordered monoid* $(M, \leqslant)$ is a monoid M equipped with a partial order $\leqslant$ which is compatible with multiplication in M; that is, $x \leqslant x'$ and $y \leqslant y'$ implies $xy \leqslant x'y'$. Every monoid can be considered as an ordered monoid by using the identity relation as order. If no ambiguity arises, we subsequently use the notation M without explicitly mentioning the order. An *order ideal* of M is a subset $I \subseteq M$ such that $y \leqslant x$ and $x \in I$ implies $y \in I$.

A *monotone homomorphism* $h \colon M \to N$ is a monoid homomorphism of ordered monoids M and N such that $x \leqslant y$ implies $h(x) \leqslant h(y)$. Submonoids of ordered monoids naturally inherit the order. A monoid N *divides* a monoid M if

there exists a surjective homomorphism from a submonoid of M onto N; moreover, if M and N are ordered, then we require the homomorphism to be monotone. The *direct product* of ordered monoids $M_1, \ldots, M_k$ is the usual direct product $M_1 \times \cdots \times M_k$ equipped with the *product order*, i.e., $(x_1, \ldots, x_k) \leqslant (y_1, \ldots, y_k)$ if $x_i \leqslant y_i$ for all $i \in \{1, \ldots, k\}$. The empty direct product is the trivial monoid.

Varieties and identities. A *variety* (respectively, *positive variety*) is a class of finite monoids (respectively, finite ordered monoids) closed under division and finite direct products. By abuse of notation, we sometimes say that an ordered monoid $(M, \leqslant)$ belongs to a variety $\mathbf{V}$ of unordered monoids if $M \in \mathbf{V}$. Both varieties and positive varieties are often defined by identities of ω-terms. We only describe the formal setting for positive varieties. The *ω-terms* over the variables X are defined inductively: The constant $1 \notin X$ is an ω-term and every variable $x \in X$ is an ω-term. If u and v are ω-terms, then so are uv and u^ω. Here, ω is considered as a formal symbol instead of a fixed integer. Every mapping $h\colon X \to M$ to a finite monoid M uniquely extends to ω-terms by setting $h(1) = 1$, $h(uv) = h(u)h(v)$ and $h(u^\omega) = h(u)^{\omega_M}$. An ordered monoid M *satisfies* the identity $U \leqslant V$ for ω-terms U and V if $h(U) \leqslant h(V)$ for all mappings $h\colon X \to M$. It satisfies $U = V$ if it satisfies both $U \leqslant V$ and $V \leqslant U$. Every class of ordered monoids defined by a set of identities of ω-terms forms a positive variety. In this paper, we need the following varieties:

- The variety $\mathbf{J}$ is the class of all so-called $\mathcal{J}$-trivial finite monoids. There are several well-known characterizations of this class, the most popular being Simon's Theorem on piecewise testable languages [23]. One can define $\mathbf{J}$ by the identities $(xyz)^\omega y = (xyz)^\omega = y(xyz)^\omega$.
- The positive variety $\mathbf{J}^+$ is defined by the identity $x \leqslant 1$. There is a language theoretic characterization similar to Simon's Theorem in terms of so-called shuffle ideals [18].
- The variety $\mathbf{DA}$ is defined by $(xyz)^\omega y(xyz)^\omega = (xyz)^\omega$. Suppose $M \in \mathbf{DA}$ and let $u, v, a \in M$. If $v \mathrel{\mathcal{R}} u \mathrel{\mathcal{R}} ua$, then $v \mathrel{\mathcal{R}} va$; and symmetrically, if $v \mathrel{\mathcal{L}} u \mathrel{\mathcal{L}} au$, then $v \mathrel{\mathcal{L}} av$, see *e.g.* [12, Lemma 1].

Languages and syntactic monoids. A language $L \subseteq A^*$ is *recognized* by a homomorphism $h\colon A^* \to M$ to some ordered monoid M if $L = h^{-1}(I)$ for some order ideal I of M. An ordered monoid M recognizes a language $L \subseteq A^*$ if there exists a homomorphism $h\colon A^* \to M$ which recognizes L. The *syntactic preorder* $\leqslant_L$ on words is defined as follows: We set $u \leqslant_L v$ for $u, v \in A^*$ if $pvq \in L$ implies $puq \in L$ for all $p, q \in A^*$. We write $u \equiv_L v$ if both $u \leqslant_L v$ and $v \leqslant_L u$. The *syntactic monoid* M_L of L is the quotient $A^*/{\equiv_L}$ consisting of the equivalence classes of $\equiv_L$; it is the unique minimal recognizer of L and it is effectively computable from any reasonable presentation of a given regular language. The syntactic preorder induces a partial order on the $\equiv_L$-classes such that M_L becomes an ordered monoid. The *syntactic homomorphism* $h_L\colon A^* \to M_L$ is the natural quotient map.

3 Two-Sided Semidirect Products of Ordered Monoids

The two-sided semidirect product of finite monoids is a useful tool for studying decompositions and hierarchies of varieties, see *e.g.* [25]. In this section, we partially extend the definition to ordered monoids. Let M be an ordered monoid and let N be a monoid. We write the operation in M additively to improve readability, which does not mean that M is commutative. A *left action* of N on M is a mapping $(n, m) \mapsto n \cdot m$ from $N \times M$ to M such that for all $m, m_1, m_2 \in M$ and all $n, n_1, n_2 \in N$ the following axioms hold:

$$n \cdot (m_1 + m_2) = n \cdot m_1 + n \cdot m_2$$
$$(n_1 n_2) \cdot m = n_1 \cdot (n_2 \cdot m)$$
$$1 \cdot m = m$$
$$n \cdot 0 = 0$$
$$n \cdot m_1 \leqslant n \cdot m_2 \quad \text{whenever} \quad m_1 \leqslant m_2$$

To shorten notation, we usually write nm instead of $n \cdot m$. A *right action* of N on M is defined symmetrically. A left and a right action are *compatible* if $(n_1 m) n_2 = n_1 (m n_2)$ for all $m \in M$ and all $n_1, n_2 \in N$. For compatible left and right actions of N on M we define the *two-sided semidirect product* $M \ast\ast N$ as the ordered monoid on the set $M \times N$ with the multiplication

$$(m_1, n_1)(m_2, n_2) = (m_1 n_2 + n_1 m_2, n_1 n_2),$$

and the order given by

$$(m_1, n_1) \leqslant (m_2, n_2) \quad \text{if and only if} \quad m_1 \leqslant m_2 \text{ and } n_1 = n_2.$$

It is straightforward to verify that $M \ast\ast N$ indeed is an ordered monoid for each pair of compatible actions. The two-sided semidirect product with left action $(n, m) \mapsto m$ and right action $(m, n) \mapsto m$ yields the direct product of M and N. In this sense the two-sided semidirect product generalizes the usual direct product.

We now define the so-called *block product* as a particular two-sided semidirect product. Let $M^{N \times N}$ be the ordered monoid of all functions from $N \times N$ to the ordered monoid M with componentwise operation. These functions are ordered by $f_1 \leqslant f_2$ if $f_1(n_1, n_2) \leqslant f_2(n_1, n_2)$ for all $n_1, n_2 \in N$. One can view $M^{N \times N}$ as the direct product of $|N|^2$ copies of M. The *block product* $M \square N$ is the two-sided semidirect product $M^{N \times N} \ast\ast N$ induced by the following pair of left and right actions. For $f \in M^{N \times N}$ and $n, n_1, n_2 \in N$ let

$$(nf)(n_1, n_2) = f(n_1, nn_2) \quad \text{and} \quad (fn)(n_1, n_2) = f(n_1 n, n_2).$$

The relationship between two-sided semidirect products and block products is the same as in the unordered case; see *e.g.* [27].

Proposition 1. *Let M, M', N, N' be monoids and suppose that M and M' are ordered. The following properties hold:*

1. *Both M and $(N, =)$ divide every two-sided semidirect product $M \ast\ast N$.*
2. *Every two-sided semidirect product $M \ast\ast N$ divides $M \square N$.*
3. *If M divides M' and N divides N', then $M \square N$ divides $M' \square N'$.*

We now extend the notion of two-sided semidirect products to varieties. For a positive variety $\mathbf{V}$ and a variety $\mathbf{W}$ we let $\mathbf{V} ** \mathbf{W}$ consist of all ordered monoids dividing a two-sided semidirect product $M ** N$ for some $M \in \mathbf{V}$ and $N \in \mathbf{W}$. For two-sided semidirect products $M ** N$ and $M' ** N'$, we define a new two-sided semidirect product $(M \times M') ** (N \times N')$ by the actions

$$(n, n')(m, m') = (nm, n'm')$$
$$(m, m')(n, n') = (mn, m'n')$$

for all $m \in M$, $m' \in M'$, $n \in N$, and $n' \in N'$. An elementary verification shows that this two-sided semidirect product is isomorphic to $(M ** N) \times (M' ** N')$, and $\mathbf{V} ** \mathbf{W}$ forms a positive variety. By Proposition 1 we see that $\mathbf{V} ** \mathbf{W}$ is identical to the positive variety generated by all block products $M \square N$ with $M \in \mathbf{V}$ and $N \in \mathbf{W}$.

For a homomorphism $h_N : A^* \to N$ we consider the alphabet $A_N = N \times A \times N$ and the length-preserving mapping $\sigma_{h_N} : A^* \to A_N^*$ defined by $\sigma_{h_N}(a_1 \cdots a_n) = b_1 \cdots b_n$, where

$$b_i = (h_N(a_1 \cdots a_{i-1}), a_i, h_N(a_{i+1} \cdots a_n))$$

for all $i \in \{1, \ldots, n\}$. The following proposition uses such mappings to characterize the languages recognized by two-sided semidirect products. It is known as the *block product principle*.

Proposition 2. *Let $\mathbf{V}$ be a positive variety, let $\mathbf{W}$ be a variety, and let $L \subseteq A^*$. The following conditions are equivalent.*

1. *L is recognized by an ordered monoid in $\mathbf{V} ** \mathbf{W}$.*
2. *There exists a homomorphism $h_N : A^* \to N$ with $N \in \mathbf{W}$ such that L is a finite union of languages of the form $\sigma_{h_N}^{-1}(L_K) \cap L_N$ with $L_K \subseteq A_N^*$ being recognized by a monoid in $\mathbf{V}$ and $L_N \subseteq A^*$ being recognized by h_N.*

4 Decidability of Negation Nesting in FO²

In this section we give two algebraic characterizations of the languages definable in the fragment Σ_m^2 of two-variable first-order logic with a restricted number of nested negations. The first description is in terms of (weakly) iterated two-sided semidirect products with $\mathcal{J}$-trivial monoids. For this we define a sequence of positive varieties by setting $\mathbf{W}_1 = \mathbf{J}^+$ and $\mathbf{W}_m = \mathbf{W}_{m-1} ** \mathbf{J}$. As for the second characterization, we define sequences of ω-terms U_m and V_m by setting

$$U_1 = z, \qquad U_m = (U_{m-1} x_m)^\omega U_{m-1} (y_m U_{m-1})^\omega,$$
$$V_1 = 1, \qquad V_m = (U_{m-1} x_m)^\omega V_{m-1} (y_m U_{m-1})^\omega,$$

where $x_2, y_2, \ldots, x_m, y_m, z$ are variables.

Theorem 1. *Let $L \subseteq A^*$ and let $m \geqslant 1$. The following conditions are equivalent:*

1. *L is definable in Σ_m^2.*
2. *The ordered syntactic monoid of L is in $\mathbf{W}_m$.*
3. *The ordered syntactic monoid of L is in $\mathbf{DA}$ and satisfies $U_m \leqslant V_m$.*

Since condition 3. in Theorem 1 is decidable for any given regular language L, this immediately yields the following corollary.

Corollary 1. *It is decidable whether a given regular language is definable in Σ_m^2.*

Note that in condition 3. of Theorem 1 one cannot drop requiring that the syntactic monoid is in **DA**. For example, the syntactic monoid of $A^* \setminus A^* aa A^*$ over $A = \{a, b\}$ satisfies the identity $U_m \leqslant V_m$ for all $m \geqslant 2$. It is nonetheless not Σ_m^2-definable, because it is not even FO^2-definable (and thus its syntactic monoid is not in **DA**). The remainder of this section proves Theorem 1. We begin with the direction $(1) \Rightarrow (2)$. The arguments are similar to Straubing's for characterizing FO_m^2 in terms of unordered two-sided semidirect products [26].

Lemma 1. *Let $m \geqslant 1$. If L is definable in Σ_m^2, then $M_L \in \mathbf{W}_m$.*

Proof. Let φ be a sentence in Σ_m^2 such that $L = L(\varphi)$. We may assume that quantifier-free subformulas of φ do not contain negations.

The proof proceeds by induction on m. For the base case $m = 1$, the language L is a finite union of languages of the form $A^* a_1 \cdots A^* a_k A^*$ and thus $pq \in L$ implies $puq \in L$ for all $p, u, q \in A^*$. This means that M_L satisfies $x \leqslant 1$ and therefore, $M_L \in \mathbf{J}^+$, see [18].

Let now $m \geqslant 2$. An *innermost block* of φ is a maximal negation-free subformula $\psi(x)$ of φ. As in the unordered case, one can show that each block is equivalent to a disjunction of formulas of the form

$$\lambda(x) = a \wedge \left(\exists y_1 \cdots \exists y_r \bigwedge_{i=1}^{r} (y_i < x \wedge \lambda(y_i) = a_i) \wedge \pi(y_1, \ldots, y_r) \right) \wedge$$
$$\left(\exists z_1 \cdots \exists z_s \bigwedge_{i=1}^{s} (z_i > x \wedge \lambda(z_i) = a_i') \wedge \pi'(z_1, \ldots, z_s) \right),$$

where π and π' are quantifier-free formulas defining an order on their parameters. Hence, each innermost block $\psi(x)$ requires that x is an a-position and that certain subwords appear to the left and to the right of position x. Let k be the maximum of all r and s occurring in these blocks. By Simon's Theorem [23], there exists an unordered monoid $N \in \mathbf{J}$ and a homomorphism $h_N \colon A^* \to N$ such that $h_N(u) = h_N(v)$ if and only if u and v agree on subwords of length at most k. Now, the aforementioned blocks can be replaced by a disjunction of formulas $\lambda(x) = (n, a, n')$ with $n, n' \in N$ and $a \in A$ to obtain an equivalent formula over the alphabet A_N.

After replacing each innermost block, the resulting formula φ' is in Σ_{m-1}^2. By induction, the corresponding language $L(\varphi')$ is recognized by a monoid $K \in \mathbf{W}_{m-1}$. We have $L = L(\varphi) = \sigma_{h_N}(L(\varphi'))$ by construction. Proposition 2 finally yields $M_L \in \mathbf{W}_{m-1} ** \mathbf{J} = \mathbf{W}_m$. $\square$

The following lemma can be seen by a similar reasoning as in the unordered case due to Straubing [26].

Lemma 2. *Let $m \geqslant 1$. If $M \in \mathbf{W}_m$, then $M \in \mathbf{DA}$ and M satisfies $U_m \leqslant V_m$.*

We turn to the implication $(3) \Rightarrow (1)$ in Theorem 1, from $U_m \leqslant V_m$ back to logic Σ_m^2. On a high-level perspective, we want to use induction on m, then use the identity $U_{m-1} \leqslant V_{m-1}$ to get to Σ_{m-1}^2, and finally lift this back to Σ_m^2. An important part of this argument is the ability to restrict (or *relativize*) the interpretation of Σ_m^2-formulas to certain factors of the model which are given by first and last occurrences of letters.

In the following we also have to take the *quantifier depth* of a formula into account, *i.e.*, the maximal number of nested quantifiers. For an integer $n \geqslant 0$ let $\Sigma_{m,n}^2$ be the fragment of Σ_m^2 of formulas with quantifier depth at most n.

Lemma 3. *Let $\varphi \in \Sigma_{m,n}^2$ for $m,n \geqslant 0$, and let $a \in A$. There exist formulas $\langle \varphi \rangle_{>\mathsf{X}a} \in \Sigma_{m,n+1}^2$ and $\langle \varphi \rangle_{<\mathsf{X}a} \in \Sigma_{m+1,n+1}^2$ such that for all $u = u_1 a u_2$ with $a \notin \mathrm{alph}(u_1)$ and $i = |u_1 a|$ we have:*

$$u,p,q \models \langle \varphi \rangle_{<\mathsf{X}a} \ \text{ if and only if } \ u_1,p,q \models \varphi \ \text{ for all } 1 \leqslant p,q < i,$$

$$u,p,q \models \langle \varphi \rangle_{>\mathsf{X}a} \ \text{ if and only if } \ u_2,p-i,q-i \models \varphi \ \text{ for all } i < p,q \leqslant |u|.$$

Proof. Let $\langle \varphi \rangle_{<\mathsf{X}a} \equiv \varphi$ if φ is an atomic formula. For conjunction and disjunction, and negation we inductively take $\langle \varphi \rangle_{<\mathsf{X}a} \wedge \langle \psi \rangle_{<\mathsf{X}a}$ and $\langle \varphi \rangle_{<\mathsf{X}a} \vee \langle \psi \rangle_{<\mathsf{X}a}$, and $\neg \langle \varphi \rangle_{<\mathsf{X}a}$, respectively. For existential quantification let

$$\langle \exists x \, \varphi \rangle_{<\mathsf{X}a} \ \equiv \ \exists x \left(\neg (\exists y \leqslant x \colon \lambda(y) = a) \wedge \langle \varphi \rangle_{<\mathsf{X}a} \right).$$

As usual, swapping the variables x and y yields the corresponding constructions for y. Atomic formulas and Boolean combinations in the construction of $\langle \varphi \rangle_{>\mathsf{X}a}$ are as above. For existential quantification let

$$\langle \exists x \, \varphi \rangle_{>\mathsf{X}a} \ \equiv \ \exists x \left((\exists y < x \colon \lambda(y) = a) \wedge \langle \varphi \rangle_{>\mathsf{X}a} \right). \qquad \square$$

The notation in the indices of the formulas mean that we restrict to the positions smaller (respectively, greater) than the first a-position (the neXt a-position, thence X_a). Of course there are dual formulas $\langle \varphi \rangle_{<\mathsf{Y}b} \in \Sigma_{m,n+1}^2$ as well as $\langle \varphi \rangle_{>\mathsf{Y}b} \in \Sigma_{m+1,n+1}^2$ for the last b-position (*i.e.*, the Yesterday b-position). The next lemma handles the case of the first a-position lying beyond the last b-position.

Lemma 4. *Let $\varphi \in \Sigma_{m,n}^2$ for $m,n \geqslant 0$, and let $a,b \in A$. There exists a formula $\langle \varphi \rangle_{(\mathsf{Y}b;\mathsf{X}a)}$ in $\Sigma_{m+1,n+1}^2$ such that for all words $u = u_1 b u_2 a u_3$ with $b \notin \mathrm{alph}(u_2 a u_3)$ and $a \notin \mathrm{alph}(u_1 b u_2)$ and for all $|u_1 b| < p,q \leqslant |u_1 b u_2|$ we have:*

$$u,p,q \models \langle \varphi \rangle_{(\mathsf{Y}b;\mathsf{X}a)} \ \text{ if and only if } \ u_2, p-|u_1 b|, q-|u_1 b| \models \varphi.$$

Proof. Atomic formulas and Boolean combinations are straightforward. Let the macro $\mathsf{Y}_b < x < \mathsf{X}_a$ stand for $\neg(\exists y \leqslant x \colon \lambda(y) = a) \wedge \neg(\exists y \geqslant x \colon \lambda(y) = b)$. Using this shortcut, we set $\langle \exists x \, \varphi \rangle_{(\mathsf{Y}b;\mathsf{X}a)} \equiv \exists x \left((\mathsf{Y}_b < x < \mathsf{X}_a) \wedge \langle \varphi \rangle_{(\mathsf{Y}b;\mathsf{X}a)} \right).$ $\qquad \square$

Let $h \colon A^* \to M$ be a homomorphism. The *$\mathcal{L}$-factorization* of a word u is the unique factorization $u = s_0 a_1 \cdots s_{\ell-1} a_\ell s_\ell$ with $s_i \in A^*$ and so-called *markers* $a_i \in A$ such that $h(s_\ell) \ \mathcal{L} \ 1$ and $h(s_i a_{i+1} \cdots s_{\ell-1} a_\ell s_\ell) >_{\mathcal{L}} h(a_i s_i \cdots a_\ell s_\ell) \ \mathcal{L} \ h(s_{i-1} a_i \cdots s_{\ell-1} a_\ell s_\ell)$ for all i. Note that $\ell < |M|$. Furthermore, if $M \in \mathbf{DA}$, then $a_i \notin \mathrm{alph}(s_i)$. Let $D_{\mathcal{L}}(u)$ consist of the positions of the markers, *i.e.*, let

$D_{\mathcal{L}}(u) = \{|s_0 a_1 \cdots s_{i-1} a_i| \mid 1 \leqslant i \leqslant \ell\}$. The $\mathcal{R}$-factorization is defined left-right symmetrically, and the set $D_{\mathcal{R}}(u)$ consists of all positions $|pa|$ for prefixes pa of u such that $h(p) >_{\mathcal{R}} h(pa)$ for some $a \in A$. The following lemma combines the $\mathcal{R}$-factorization with the $\mathcal{L}$-factorization for monoids in $\mathbf{DA}$ such that, starting with Σ_m^2, one can express Σ_{m-1}^2-properties of the factors. To formulate this feature we set $u \leqslant_{m,n}^{\mathrm{FO}^2} v$ for words $u, v \in A^*$ if $v \models \varphi$ implies $u \models \varphi$ for all $\varphi \in \Sigma_{m,n}^2$.

Lemma 5. *Let $h\colon A^* \to M$ be a homomorphism with $M \in \mathbf{DA}$, let $m \geqslant 2$ and $n \geqslant 0$ be integers, and let $u, v \in A^*$ with $u \leqslant_{m,2|M|+n}^{\mathrm{FO}^2} v$. There exist factorizations $u = s_0 a_1 \cdots s_{\ell-1} a_\ell s_\ell$ and $v = t_0 a_1 \cdots t_{\ell-1} a_\ell t_\ell$ with $a_i \in A$ and $s_i, t_i \in A^*$ such that the following properties hold for all $i \in \{1, \ldots, \ell\}$:*

1. *$s_i \leqslant_{m-1,n}^{\mathrm{FO}^2} t_i$,*
2. *$h(s_0) \; \mathcal{R} \; 1$ and $h(t_0 a_1 \cdots t_{i-1} a_i) \; \mathcal{R} \; h(t_0 a_1 \cdots t_{i-1} a_i s_i)$,*
3. *$h(s_\ell) \; \mathcal{L} \; 1$ and $h(a_i s_i \cdots a_\ell s_\ell) \; \mathcal{L} \; h(s_{i-1} a_i \cdots a_\ell s_\ell)$.*

Proof. Note that in property 2. the suffix is s_i and not t_i. We want to prove the claim by an induction, for which we have to slightly generalize the claim. Apart from the words u and v from the premises of the lemma we also consider an additional word p which serves as a prefix for v. The proof is by induction on $|D_{\mathcal{R}}(pv) \setminus D_{\mathcal{R}}(p)|$. The assumptions are $u \leqslant_{m,n'}^{\mathrm{FO}^2} v$, where $n' = n + |D_{\mathcal{R}}(pv) \setminus D_{\mathcal{R}}(p)| + |D_{\mathcal{L}}(u)| + 1$. We shall construct factorizations $u = s_0 a_1 \cdots s_{\ell-1} a_\ell s_\ell$ and $pv = p t_0 a_1 \cdots t_{\ell-1} a_\ell t_\ell$ such that properties 1. and 3. hold, but instead of 2. we have $h(p t_0 a_1 \cdots t_{i-1} a_i) \; \mathcal{R} \; h(p t_0 a_1 \cdots t_{i-1} a_i s_i)$ and $h(p s_0) \; \mathcal{R} \; h(p)$. We thus recover the lemma using an empty prefix p.

Let $u = s_0' c_1 \cdots s_{\ell'-1}' c_{\ell'} s_{\ell'}'$ be the $\mathcal{L}$-factorization (in particular $c_i \notin \mathrm{alph}(s_i')$) and let $v = t_0' c_1 \cdots t_{\ell'-1}' c_{\ell'} t_{\ell'}'$ where $c_i \notin \mathrm{alph}(t_i')$ for all i. The factorization of v exists because by assumption u and v agree on subwords of length ℓ'. The dual of Lemma 3 yields $s_0' c_1 \cdots c_{\ell'-i} s_{\ell'-i}' \leqslant_{m-1,n'-i}^{\mathrm{FO}^2} t_0' c_1 \cdots c_{\ell'-i} t_{\ell'-i}'$ as well as $s_i' \leqslant_{m-1,n}^{\mathrm{FO}^2} t_i'$ for all i.

First suppose $D_{\mathcal{R}}(p) = D_{\mathcal{R}}(pv)$. In this case $h(p) \; \mathcal{R} \; h(pv)$, and therefore, $h(p) \; \mathcal{R} \; h(px)$ for all $x \in B^*$, where $B = \mathrm{alph}(v)$. So in particular we have that $h(p t_0' c_1 \cdots t_{i-1}' c_i) \; \mathcal{R} \; h(p t_0' c_1 \cdots t_{i-1}' c_i s_i')$ because $\mathrm{alph}(u) = B$. Setting $a_i = c_i$, $s_i = s_i'$, and $t_i = t_i'$ yields a factorization with the desired properties.

Suppose now $D_{\mathcal{R}}(p) \subsetneq D_{\mathcal{R}}(pv)$, and let s be the longest prefix of u such that $h(p) \; \mathcal{R} \; h(ps) >_{\mathcal{R}} h(psa)$ for some $a \in A$. Such a prefix exists as $\mathrm{alph}(u) = \mathrm{alph}(v)$. We have $a \notin \mathrm{alph}(s)$ by $M \in \mathbf{DA}$. Let t be the longest prefix of v with $a \notin \mathrm{alph}(t)$. Using Lemma 3 we see $\mathrm{alph}(t) \subseteq \mathrm{alph}(s)$. Let k and k' be maximal such that $s_0' c_1 \cdots s_{k-1}' c_k$ is a prefix of s and such that $t_0' c_1 \cdots t_{k'-1}' c_{k'}$ is a prefix of t. We claim $k = k'$. For instance, suppose $k < k'$. Then $a c_{k+1} \cdots c_{\ell'}$ is a subword of u but not of v (since $c_{k+1} t_{k+1}' \cdots c_{\ell'} t_{\ell'}'$ is the shortest suffix of v with the subword $c_{k+1} \cdots c_{\ell'}$ and since there is no a-position in $t_0' c_1 \cdots t_k'$). Let $a_i = c_i$ for $i \in \{1, \ldots, k\}$, let $s_i = s_i'$ and $t_i = t_i'$ for $i \in \{0, \ldots, k-1\}$. Let s_k and t_k such that $s = s_0 c_1 \cdots s_{k-1} c_k s_k$ and $t = t_0 c_1 \cdots t_{k-1} c_k t_k$. Lemma 4 yields $s_k \leqslant_{m-1,n}^{\mathrm{FO}^2} t_k$.

Let $u = sau'$ and $v = tav'$, and let $p' = pta$. For all $i \in \{0, \ldots, k\}$ we have $h(p t_0 a_1 \cdots t_{i-1} a_i) \; \mathcal{R} \; h(p t_0 a_1 \cdots t_{i-1} a_i s_i)$ because $\mathrm{alph}(t) \subseteq \mathrm{alph}(s)$. Note

that $h(a_{i+1}s_{i+1}\cdots a_k s_k a u') \mathcal{L}\ h(s_i a_{i+1} s_{i+1} \cdots a_k s_k a u')$. Since $M \in \mathbf{DA}$ we see $h(p) >_{\mathcal{R}} h(p')$ and thus $D_{\mathcal{R}}(p) \subsetneq D_{\mathcal{R}}(p')$. Using the formulas $\langle \varphi \rangle_{>\mathsf{X}a}$ from Lemma 3 yields $u' \leqslant_{m,n'-1}^{\mathrm{FO}^2} v'$. As $n' \geqslant |D_{\mathcal{R}}(p'v') \setminus D_{\mathcal{R}}(p')| + |D_{\mathcal{L}}(u')| + 2$ we can apply induction to obtain factorizations $u' = s_{k+1} a_{k+2} \cdots s_{\ell-1} a_\ell s_\ell$ and $v' = t_{k+1} a_{k+2} \cdots t_{\ell-1} a_\ell t_\ell$. Setting $a_{k+1} = a$ yields the desired factorizations. $\quad\square$

The preceding lemma enables induction on the parameter m. We start with a homomorphism onto a monoid satisfying $U_m \leqslant V_m$ and want to show that preimages of $\leqslant$-order ideals are unions of $\leqslant_{m,n}^{\mathrm{FO}^2}$-order ideals for some sufficiently large n. Intuitively, a string rewriting technique yields the largest quotient which satisfies the identity $U_{m-1} \leqslant V_{m-1}$. One rewriting step corresponds to one application of the identity $U_{m-1} \leqslant V_{m-1}$ of level $m-1$. Such rewriting steps can be lifted to the identity $U_m \leqslant V_m$ in the contexts they are applied.

Proposition 3. *Let $m \geqslant 1$ be an integer, let $h\colon A^* \to M$ be a surjective homomorphism onto an ordered monoid $M \in \mathbf{DA}$ satisfying $U_m \leqslant V_m$. There exists a positive integer n such that $u \leqslant_{m,n}^{\mathrm{FO}^2} v$ implies $h(u) \leqslant h(v)$ for all $u, v \in A^*$.*

Proof. We proceed by induction on m. For the base case $m = 1$ a result of Pin [18] shows that, for every $\leqslant$-order ideal I of M, the set $h^{-1}(I)$ is a finite union of languages $A^* a_1 \cdots A^* a_k A^*$ for some $k \geqslant 1$ and $a_i \in A$. Let n be the maximum of all indices k appearing in those unions when considering all order ideals $I \subseteq M$. If $u \leqslant_{1,n}^{\mathrm{FO}^2} v$, then for all languages $P = A^* a_1 \cdots A^* a_k A^*$ with $k \leqslant n$ we have that $v \in P$ implies $u \in P$. Moreover, the preimage L of the order ideal generated by $h(v)$ is a finite union of languages $A^* a_1 \cdots A^* a_k A^*$ with $k \leqslant n$. We have $v \in L$ and thus $u \in L$. This shows $h(u) \leqslant h(v)$.

In the following let $m \geqslant 2$ and fix some integer $\omega \geqslant 1$ such that x^ω is idempotent for all $x \in M$. We introduce a string rewriting system $\to$ on A^* by letting $t \to s$ if $h(s) = h(t)$ or if $t = p v_{m-1} q$ and $s = p u_{m-1} q$ for $p, q \in A^*$, and $v_1 = 1$ and $u_1 = z$, and for $i \geqslant 2$ we have

$$v_i = (u_{i-1} x_i)^\omega v_{i-1} (y_i u_{i-1})^\omega, \quad u_i = (u_{i-1} x_i)^\omega u_{i-1} (y_i u_{i-1})^\omega$$

for $x_i, y_i, z \in A^*$. Note that $t \to s$ implies $p' t q' \to p' s q'$ for all $p', q' \in A^*$. Let $\overset{*}{\to}$ be the transitive closure of $\to$, *i.e.*, let $t \overset{*}{\to} s$ if there exists a chain $t = w_1 \to w_2 \to \cdots \to w_\ell = s$ of rewriting steps for some $\ell \geqslant 1$ and $w_i \in A^*$. We claim that we can lift the rewriting steps of $t \overset{*}{\to} s$ to M within certain contexts in an order respecting way.

Claim. Let $u, v, s, t \in A^*$ with $t \overset{*}{\to} s$. If both $h(u) \mathcal{R}\ h(us)$ and $h(v) \mathcal{L}\ h(sv)$, then $h(usv) \leqslant h(utv)$.

The proof of the claim is by induction on the length of a minimal $\to$-chain from t to s. The claim is trivial if $h(t) = h(s)$. Suppose $t \overset{*}{\to} t' \to s$ and $t' = p v_{m-1} q$ and $s = p u_{m-1} q$. Since $h(u) \mathcal{R}\ h(us)$, there exists $x \in A^*$ such that $h(u) = h(usx)$; and since $h(v) \mathcal{L}\ h(sv)$ there exists $y \in A^*$ such that $h(v) = h(ysv)$. Now $h(u) = h\big(u(pu_{m-1}qx)^\omega\big)$ and $h(v) = h\big((ypu_{m-1}q)^\omega v\big)$. By letting $x_m = qxp$ and $y_m = qyp$, the identity $U_m \leqslant V_m$ of M yields

$$h(usv) = h\big(up(u_{m-1}x_m)^\omega u_{m-1}(y_m u_{m-1})^\omega qv\big)$$
$$\leqslant h\big(up(u_{m-1}x_m)^\omega v_{m-1}(y_m u_{m-1})^\omega qv\big) = h(ut'v).$$

Observe that $(pu_{m-1}qx)^\omega p = p(u_{m-1}qxp)^\omega = p(u_{m-1}x_m)^\omega$. Note that $\mathrm{alph}(t') \subseteq \mathrm{alph}(s)$. Therefore, $h(u) \ \mathcal{R} \ h(us)$ implies $h(u) \ \mathcal{R} \ h(ut')$, and symmetrically $h(v) \ \mathcal{L} \ h(sv)$ implies $h(v) \ \mathcal{L} \ h(t'v)$. Induction yields $h(ut'v) \leqslant h(utv)$ and thus $h(usv) \leqslant h(utv)$. This completes the proof of the claim.

Let $t \sim s$ if $t \overset{*}{\to} s$ and $s \overset{*}{\to} t$. Let M' be the quotient $A^*/{\sim}$. The relation $\sim$ is a congruence on A^* and M' is naturally equipped with a monoid structure. Let $h' \colon A^* \to M'$ be the canonical homomorphism mapping $u \in A^*$ to its equivalence class modulo $\sim$. The preorder $\overset{*}{\to}$ on A^* induces a partial order on M' by letting $h'(u) \leqslant h'(v)$ whenever $v \overset{*}{\to} u$. Thus M' forms an ordered monoid. Moreover, M' is an unordered quotient of M and, in particular, M' is finite and in $\mathbf{DA}$, and x^ω is idempotent for all $x \in M'$.

By construction, M' satisfies the identity $U_{m-1} \leqslant V_{m-1}$ and induction yields an integer n such that $u \leqslant^{\mathrm{FO}^2}_{m-1,n} v$ implies $h'(u) \leqslant h'(v)$. We show that $u \leqslant^{\mathrm{FO}^2}_{m,n'} v$ implies $h(u) \leqslant h(v)$ for $n' = n + 2|M|$. Suppose $u \leqslant^{\mathrm{FO}^2}_{m,n'} v$ and consider the factorizations $u = s_0 a_1 \cdots s_{\ell-1} a_\ell s_\ell$ and $v = t_0 a_1 \cdots t_{\ell-1} a_\ell t_\ell$ from Lemma 5. For all i we have:

 - $s_i \leqslant^{\mathrm{FO}^2}_{m-1,n} t_i$ and thus $t_i \overset{*}{\to} s_i$ by choice of n,
 - $h(t_0 a_1 \cdots t_{i-1} a_i) \ \mathcal{R} \ h(t_0 a_1 \cdots t_{i-1} a_i s_i)$, and
 - $h(a_{i+1} s_{i+1} \cdots a_\ell s_\ell) \ \mathcal{L} \ h(s_i a_{i+1} s_{i+1} \cdots a_\ell s_\ell)$.

For conciseness $t_0 a_1 \cdots t_{i-1} a_i$ is the empty word if $i = 0$, and so is $a_{i+1} s_{i+1} \cdots a_\ell s_\ell$ if $i = \ell$. Applying the above claim repeatedly to substitute s_i with t_i for increasing $i \in \{0, \ldots, \ell\}$ yields the following chain of inequalities:

$$\begin{aligned}
h(u) = h(s_0 a_1 s_1 \cdots s_{\ell-1} a_\ell s_\ell) \\
\leqslant h(t_0 a_1 s_1 \cdots s_{\ell-1} a_\ell s_\ell) \\
\vdots \\
\leqslant h(t_0 a_1 t_1 \cdots t_{\ell-1} a_\ell s_\ell) \\
\leqslant h(t_0 a_1 t_1 \cdots t_{\ell-1} a_\ell t_\ell) = h(v). \qquad \square
\end{aligned}$$

Proof of Theorem 1. The implication 1. $\Rightarrow$ 2. is Lemma 1, and 2. $\Rightarrow$ 3. is Lemma 2. For the implication 3. $\Rightarrow$ 1., let $L \subseteq A^*$ be a language, let $h_L \colon A^* \to M_L$ be its syntactic homomorphism. Moreover, suppose that M_L is in $\mathbf{DA}$ and satisfies $U_m \leqslant V_m$. The set $I = h_L(L)$ is an order ideal of M_L. Proposition 3 shows that there exists an integer n such that $L = h_L^{-1}(I)$ is a union of $\leqslant^{\mathrm{FO}^2}_{m,n}$-order ideals. Up to equivalence, there are only finitely many formulas with quantifier depth n. Therefore, $\leqslant^{\mathrm{FO}^2}_{m,n}$-order ideals are $\Sigma^2_{m,n}$-definable. $\qquad \square$

Conclusion

The fragments Σ^2_m of $\mathrm{FO}^2[<]$ are defined by restricting the number of nested negations. They can be seen as the half levels of the alternation hierarchy FO^2_m in two-variable first-order logic, and we have $\Sigma^2_m \subseteq \mathrm{FO}^2_m \subseteq \Sigma^2_{m+1}$. It is known that the languages definable in FO^2_m form a strict hierarchy, see *e.g.* [16]. For every $m \geqslant 1$ we have given ω-terms U_m and V_m such that a language L is definable in

Σ_m^2 if and only if its ordered syntactic monoid is in the variety **DA** and satisfies the identity $U_m \leqslant V_m$. Using this characterization one can decide whether a given regular language is definable in Σ_m^2. In particular, we have shown decidability for every level of an infinite hierarchy. Note that there is no immediate connection between the decidability of FO_m^2 and the decidability of Σ_m^2.

The block product principle is an important tool in the proof of the direction from Σ_m^2 to identities. In order to be able to apply this tool, we first extended block products to the case where the left factor is an ordered monoid and then stated the block product principle in this context. In order to further extend the block product $M \square N$ to the case where both M and N are ordered, one has to consider the monotone functions in $N \times N \to M$ instead of $M^{N \times N}$. As in the case of the wreath product principle [20] this leads to ordered alphabets when stating the block product principle. However, one implication in the block product principle fails for ordered alphabets as the universal property does not hold in this setting.

References

1. J. R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.
2. R. S. Cohen and J. A. Brzozowski. Dot-depth of star-free events. *J. Comput. Syst. Sci.*, 5(1):1–16, 1971.
3. V. Diekert, P. Gastin, and M. Kufleitner. A survey on small fragments of first-order logic over finite words. *Int. J. Found. Comput. Sci.*, 19(3):513–548, 2008.
4. H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Perspectives in Mathematical Logic. Springer, 1995.
5. C. C. Elgot. Decision problems of finite automata design and related arithmetics. *Trans. Amer. Math. Soc.*, 98:21–51, 1961.
6. Ch. Glaßer and H. Schmitz. Languages of dot-depth 3/2. *Theory of Computing Systems*, 42(2):256–286, 2008.
7. J. A. W. Kamp. *Tense Logic and the Theory of Linear Order*. PhD thesis, University of California, 1968.
8. R. Knast. A semigroup characterization of dot-depth one languages. *RAIRO, Inf. Théor.*, 17(4):321–330, 1983.
9. A. Krebs and H. Straubing. An effective characterization of the alternation hierarchy in two-variable logic. In *FSTTCS 2012, Proceedings*, volume 18 of *LIPIcs*, pp. 86–98. Dagstuhl Publishing, 2012.
10. M. Kufleitner and A. Lauser. Languages of dot-depth one over infinite words. In *LICS 2011, Proceedings*, pp. 23–32. IEEE Computer Society, 2011.
11. M. Kufleitner and A. Lauser. Around dot-depth one. *Int. J. Found. Comput. Sci.*, 23(6):1323–1339, 2012.
12. M. Kufleitner and A. Lauser. The join levels of the Trotter-Weil hierarchy are decidable. In *MFCS 2012, Proceedings*, volume 7464 of *LNCS*, pp. 603–614. Springer, 2012.
13. M. Kufleitner and A. Lauser. The join of the varieties of R-trivial and L-trivial monoids via combinatorics on words. *Discrete Math. & Theor. Comput. Sci.*, 14(1):141–146, 2012.

14. M. Kufleitner and A. Lauser. Quantifier alternation in two-variable first-order logic with successor is decidable. In *STACS 2013, Proceedings*, volume 20 of *LIPIcs*, pp. 305–316. Dagstuhl Publishing, 2013.

15. M. Kufleitner and P. Weil. The FO^2 alternation hierarchy is decidable. In *CSL 2012, Proceedings*, volume 16 of *LIPIcs*, pp. 426–439. Dagstuhl Publishing, 2012.

16. M. Kufleitner and P. Weil. On logical hierarchies within FO^2-definable languages. *Log. Methods Comput. Sci.*, 8:1–30, 2012.

17. R. McNaughton and S. Papert. *Counter-Free Automata*. The MIT Press, 1971.

18. J.-É. Pin. A variety theorem without complementation. In *Russian Mathematics (Iz. VUZ)*, volume 39, pp. 80–90, 1995.

19. J.-É. Pin and P. Weil. Polynomial closure and unambiguous product. *Theory Comput. Syst.*, 30(4):383–422, 1997.

20. J.-É. Pin and P. Weil. The wreath product principle for ordered semigroups. *Commun. Algebra*, 30(12):5677–5713, 2002.

21. M. P. Schützenberger. On finite monoids having only trivial subgroups. *Inf. Control*, 8:190–194, 1965.

22. Th. Schwentick, D. Thérien, and H. Vollmer. Partially-ordered two-way automata: A new characterization of DA. In *DLT 2001, Proceedings*, volume 2295 of *LNCS*, pp. 239–250. Springer, 2002.

23. I. Simon. Piecewise testable events. In *Autom. Theor. Form. Lang., 2nd GI Conf.*, volume 33 of *LNCS*, pp. 214–222. Springer, 1975.

24. H. Straubing. A generalization of the Schützenberger product of finite monoids. *Theor. Comput. Sci.*, 13:137–150, 1981.

25. H. Straubing. *Finite Automata, Formal Logic, and Circuit Complexity*. Birkhäuser, 1994.

26. H. Straubing. Algebraic characterization of the alternation hierarchy in FO^2[<] on finite words. In *CSL 2011, Proceedings*, volume 12 of *LIPIcs*, pp. 525–537. Dagstuhl Publishing, 2011.

27. H. Straubing and D. Thérien. Weakly iterated block products of finite monoids. In *LATIN 2002, Proceedings*, volume 2286 of *LNCS*, pp. 91–104. Springer, 2002.

28. P. Tesson and D. Thérien. Diamonds are forever: The variety DA. In *Semigroups, Algorithms, Automata and Languages 2001, Proceedings*, pp. 475–500. World Scientific, 2002.

29. D. Thérien. Classification of finite monoids: The language approach. *Theor. Comput. Sci.*, 14(2):195–208, 1981.

30. D. Thérien and Th. Wilke. Over words, two variables are as powerful as one quantifier alternation. In *STOC 1998, Proceedings*, pp. 234–240. ACM Press, 1998.

31. W. Thomas. Classifying regular events in symbolic logic. *J. Comput. Syst. Sci.*, 25:360–376, 1982.

32. B. A. Trakhtenbrot. Finite automata and logic of monadic predicates (in Russian). *Dokl. Akad. Nauk SSSR*, 140:326–329, 1961.

33. Ph. Weis and N. Immerman. Structure theorem and strict alternation hierarchy for FO^2 on words. *Log. Methods Comput. Sci.*, 5:1–23, 2009.