Around Dot-Depth One *

(Extended Abstract)

Manfred Kufleitner Alexander Lauser

University of Stuttgart, FMI

{kufleitner, lauser}@fmi.uni-stuttgart.de

The dot-depth hierarchy is a classification of star-free languages. It is related to the quantifier alternation hierarchy of first-order logic over finite words. We consider fragments of languages with dot-depth 1/2 and dot-depth 1 obtained by prohibiting the specification of prefixes or suffixes. As it turns out, these language classes are in one-to-one correspondence with fragments of alternation-free first-order logic without min- or max-predicate. For all fragments, we obtain effective algebraic characterizations. Moreover, we give new proofs for the decidability of the membership problem for dot-depth 1/2 and dot-depth 1.

1 Introduction

The dot-depth hierarchy $\mathcal{B}_n$ for $n \in \mathbb{N} + \{1/2, 1\}$ has been introduced by Cohen and Brzozowski [3]. A very similar hierarchy is the Straubing-Thérien hierarchy $\mathcal{L}_n$, see [19, 21]. Both hierarchies are strict [2] and they exhaust the class of star-free languages. A classical result of McNaughton and Papert is that a language is star-free if and only if it is definable in first-order logic [10]. Thomas [23] has tightened this result by showing that there is a one-to-one correspondence between the dot-depth hierarchy (and also between the Straubing-Thérien hierarchy) and the quantifier alternation hierarchy of first-order logic. More precisely, the dot-depth hierarchy is related to the quantifier alternation hierarchy over the signature $[<, +1, \min, \max]$, whereas the Straubing-Thérien hierarchy corresponds to the quantifier alternation hierarchy over the signature $[<]$.

Schützenberger has shown that a language is star-free if and only if its syntactic semigroup is aperiodic [16]. The latter property is decidable. Together with the result of McNaughton and Papert, this yields a decision procedure for definability in first-order logic. Effectively determining the level of a language in the dot-depth hierarchy or equivalently, in the quantifier alternation hierarchy of first-order logic, is one of the most

*Supported by the German Research Foundation (DFG) under grant DI 435/5-1.

challenging open problems in automata theory. For $n \in \mathbb{N}$, Straubing has shown that membership in $\mathcal{B}_n$ is decidable if and only if membership in $\mathcal{L}_n$ is decidable [20]. This result has been extended to the half-levels by Pin and Weil [15]. Simon has shown that the class of piecewise testable languages $\mathcal{L}_1$ is decidable [17]. Later, Knast [6] gave an effective algebraic characterization of $\mathcal{B}_1$. Decidability of $\mathcal{L}_{1/2}$ was shown by Pin [12], and the levels $\mathcal{B}_{1/2}$ and $\mathcal{L}_{3/2}$ are decidable by a result of Pin and Weil [14]. The most recent decidability result is for $\mathcal{B}_{3/2}$ due to Glaßer and Schmitz [4]. To date, no other levels are known to be decidable.

In this paper, we focus on subclasses of $\mathcal{B}_{1/2}$ and $\mathcal{B}_1$. For both $\mathcal{B}_{1/2}$ and $\mathcal{B}_1$ we give new proofs for their effective algebraic characterizations. The proof of Pin and Weil [14] for $\mathcal{B}_{1/2}$ is based on factorization forests [18], and the proof of Knast [6] as well as the simplified version of Thérien [22] for $\mathcal{B}_1$ are based on a generalization of finite monoids, so-called *finite categories* [24]. Our proof for $\mathcal{B}_1$ is a generalization of Klíma's proof [5] for $\mathcal{L}_1$. The main advantage of our proofs for $\mathcal{B}_{1/2}$ and $\mathcal{B}_1$ over previous ones is that the constants involved in finding language descriptions for given algebraic objects are more explicit (and therefore smaller).

The main original contributions of this paper are effective algebraic characterizations of fragments of alternation-free first-order logic over the signatures $[<, +1, \min]$ without max-predicate, $[<, +1, \max]$ without min, and $[<, +1]$ without min and max. These fragments also admit language characterizations in terms of subclasses of $\mathcal{B}_{1/2}$ and $\mathcal{B}_1$. The corresponding language classes are obtained by prohibiting the specification of prefixes or suffixes. A more detailed overview of our results can be found in the summary in Section 7.

A full version of this paper can be found as technical report [8].

2 Preliminaries

Words and languages Let Γ be a finite nonempty alphabet. The set of finite words is Γ^*. By 1 we denote the empty word and $\Gamma^+ = \Gamma^* \setminus \{1\}$ is the set of finite nonempty words. A word $v \in \Gamma^*$ is a *prefix* (resp. *suffix*, resp. *factor*) of u if $u \in v\Gamma^*$ (resp. $u \in \Gamma^* v$, resp. $u \in \Gamma^* v \Gamma^*$). The length of a word $u \in \Gamma^*$ is $|u|$ and its alphabet is $\mathrm{alph}(u) = \{a \in \Gamma \mid u \in \Gamma^* a \Gamma^*\}$. Similarly, $\mathrm{alph}_k(u) = \left\{v \in \Gamma^k \mid u \in \Gamma^* v \Gamma^*\right\}$ is the set of all factors of u of length k. A *quotient* of $L \subseteq \Gamma^+$ is a language of the form $u^{-1}L = \{v \in \Gamma^+ \mid uv \in L\}$ or $Lu^{-1} = \{v \in \Gamma^+ \mid vu \in L\}$ for $u \in \Gamma^*$. A language L is a *monomial* of degree m if $L = w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$ for some $w_1, \ldots, w_n \in \Gamma^*$ with $|w_1 \cdots w_n| = m$. A language has *dot-depth one* if it is a Boolean combination of monomials. Throughout this paper, Boolean operations are complementation, *finite* union, and *finite* intersection. Positive Boolean operations are finite union and finite intersection.

First-order logic over words We consider the first-order logic $\mathrm{FO} = \mathrm{FO}[<, +1, \min, \max]$ over nonempty finite words. We view words as sequences of labeled positions which are linearly ordered by $<$. Variables are interpreted as positions of a word. For variables x, y we have the following atomic formulas: $x < y$ says that x is a position smaller than y; and

$x = y + 1$ is true if x is the immediate successor of y; the formula $\min(x)$ (resp. $\max(x)$) holds if x is the first (resp. last) position. Moreover, we always assume that we have an atomic formula $\top$ (for *true*), equality of positions $x = y$, and a predicate $\lambda(x) = a$ specifying that position x is labeled by $a \in \Gamma$. Formulas can be composed using Boolean operations, existential quantification, and universal quantification. Their semantics is as usual. A sentence is a formula without free variables. For a sentence φ of FO we write $u \models \varphi$ if u is a model of φ and the language defined by φ is $L(\varphi) = \{u \in \Gamma^+ \mid u \models \varphi\}$.

The fragment Σ_1 consists of all FO-formulas in prenex normal form with only one block of quantifiers and these quantifiers are existential. Let $\mathcal{C} \subseteq \{<, +1, \min, \max\}$. By $\Sigma_1[\mathcal{C}]$ we denote the class of formulas in Σ_1 which only use predicates in $\mathcal{C}$, equality, and the label predicate. The fragment of alternation-free formulas over the signature $\mathcal{C}$ is $\mathbb{B}\Sigma_1[\mathcal{C}]$; it comprises all Boolean combinations of formulas in $\Sigma_1[\mathcal{C}]$.

Finite semigroups and recognizable languages Let S be a semigroup. We always assume that S is nonempty. The set of *idempotents* is $E(S) = \{e \in S \mid e^2 = e\}$. For every finite semigroup S there exists a number $\omega \geq 1$ such that for every $x \in S$, the power x^ω is the unique idempotent element generated by x. Frequently, we consider words $u, v \in S^*$ where the alphabet is a semigroup. We write "$u = v$ in S" if either $u = 1 = v$ or $u, v \in S^+$ evaluate to the same element of S.

Lemma 1. *Let S be a finite semigroup. For all $x_1, \ldots, x_{|S|} \in S$ there exist $i \in \{1, \ldots, |S|\}$ and $e \in E(S)$ such that $x_1 \cdots x_i = x_1 \cdots x_i e$ in S.*

A subset $I \subseteq S$ is an *ideal* (resp. *right ideal*, resp. *left ideal*) if $S^1 I S^1 \subseteq I$ (resp. $IS^1 \subseteq I$, resp. $S^1 I \subseteq I$). Here, the monoid $S^1 = S \cup \{1\}$ is obtained by adjoining a neutral element. Green's relations are an important tool in the study of semigroups. They are defined as follows. Let $x \leq_{\mathcal{J}} y$ (resp. $x \leq_{\mathcal{R}} y$, resp. $x \leq_{\mathcal{L}} y$) if there exist $s, t \in S^1$ such that $x = syt$ in S (resp. $x = yt$ in S, resp. $x = sy$ in S). Let $x \mathrel{\mathcal{J}} y$ (resp. $x \mathrel{\mathcal{R}} y$, resp. $x \mathrel{\mathcal{L}} y$) if $x \leq_{\mathcal{J}} y$ and $y \leq_{\mathcal{J}} x$ (resp. $x \leq_{\mathcal{R}} y$ and $y \leq_{\mathcal{R}} x$, resp. $x \leq_{\mathcal{L}} y$ and $y \leq_{\mathcal{L}} x$). Therefore, $x \mathrel{\mathcal{J}} y$ (resp. $x \mathrel{\mathcal{R}} y$, resp. $x \mathrel{\mathcal{L}} y$) if and only if x and y generate the same ideals (resp. right ideals, resp. left ideals) in S. The relations $\leq_{\mathcal{R}}$, $\leq_{\mathcal{L}}$, and $\leq_{\mathcal{J}}$ form preorders on S; therefore $\mathcal{R}$, $\mathcal{L}$, and $\mathcal{J}$ are equivalence relations.

Let $\leq$ be a preorder on S. A set $P \subseteq S$ is a *$\leq$-order ideal* if $x \leq y \in P$ implies $x \in P$. An *ordered semigroup* S is equipped with a compatible partial order $\leq$, i.e., if $p \leq q$ and $s \leq t$, then $ps \leq qt$. Every semigroup is an ordered semigroup with equality as partial order. A language $L \subseteq \Gamma^+$ is *recognized* by an ordered semigroup S if there exists a homomorphism $h : \Gamma^+ \to S$ such that $L = h^{-1}(P)$ for some $\leq$-order ideal P. If the order of S is equality, then we obtain the usual notion of recognition. Note that every $\leq_{\mathcal{R}}$-order ideal (resp. $\leq_{\mathcal{L}}$-order ideal, resp. $\leq_{\mathcal{J}}$-order ideal) is a right ideal (resp. left ideal, resp. ideal) and vice versa. For a language $L \subseteq \Gamma^+$ the *syntactic preorder* $\leq_L$ over Γ^+ is given by $x \leq_L y$ if $uyv \in L \Rightarrow uxv \in L$ for all $u, v \in \Gamma^*$. The syntactic congruence $\equiv_L$ is defined by $x \equiv_L y$ if both $x \leq_L y$ and $y \leq_L x$. The equivalence classes $[x]_L = \{y \in \Gamma^+ \mid x \equiv_L y\}$ equipped with the canonical composition constitute the *syntactic semigroup* $\mathrm{Synt}(L)$ and the preorder $\leq_L$ of Γ^+ becomes a compatible partial order of $\mathrm{Synt}(L)$. The *syntactic* *homomorphism* is $h_L : \Gamma^+ \to \mathrm{Synt}(L)$ with $h_L(x) = [x]_L$. The syntactic semigroup of L is finite if and only if L is regular. Moreover, every language is recognized by its syntactic semigroup.

By $[\![x^\omega y x^\omega \leq x^\omega]\!]$ we denote the class of finite ordered semigroups S such that $x^\omega y x^\omega \leq x^\omega$ for all elements $x, y \in S$. We let $\mathbf{B_1}$ be the class of finite semigroups S such that $(exfy)^\omega exf(tesf)^\omega = (exfy)^\omega esf(tesf)^\omega$ for all idempotents $e, f \in E(S)$ and all elements $s, t, x, y \in S$.

Lemma 2. *Let $(S, \leq)$ be an ordered semigroup such that $x^\omega y x^\omega \leq x^\omega$ for all $x, y \in S$. Then $S \in \mathbf{B_1}$.*

Lemma 3. *Let $S \in \mathbf{B_1}$ and let $u, v \in S$ with $u = ue$ and $v = ve$ for some idempotent $e \in E(S)$. If $u \mathrel{\mathcal{R}} v$, then $u = v$.*

3 Dot-depth 1/2

A language $L \subseteq \Gamma^+$ has dot-depth 1/2 if it is a positive Boolean combination of monomials $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$ with $w_i \in \Gamma^*$. By a result of Thomas [23], a language has dot-depth 1/2 if and only if it is definable in existential first-order logic $\Sigma_1[<, +1, \min, \max]$. Pin and Weil [14] have shown that L has dot-depth 1/2 if and only if $\mathrm{Synt}(L) \in [\![x^\omega y x^\omega \leq x^\omega]\!]$. In this section, we give a new proof of these equivalences. The key step in the proof is to show that if $L \subseteq \Gamma^+$ is recognized by some semigroup in $[\![x^\omega y x^\omega \leq x^\omega]\!]$, then L is a union of monomials $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$. The main advantage of the proof given here is that the degree $|w_1 \cdots w_n|$ is polynomially bounded (Proposition 9), whereas in the proof of Pin and Weil, the bound is exponential.

Theorem 4 (Pin/Weil [14], Thomas [23])**.** *Let $L \subseteq \Gamma^+$. The following are equivalent:*

1. *L is definable in $\Sigma_1[<, +1, \min, \max]$.*
2. *L is a finite union of monomials $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$.*
3. *L is a positive Boolean combination of monomials $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$.*
4. $\mathrm{Synt}(L) \in [\![x^\omega y x^\omega \leq x^\omega]\!]$.

In the remainder of this section we prove the above theorem.

Lemma 5. *Let $L \subseteq \Gamma^+$ be definable by a sentence in $\Sigma_1[<, +1, \min, \max]$ with m variables. Then L is a finite union of languages $w_1 \Gamma^+ w_2 \cdots \Gamma^+ w_n$ with $|w_1 \cdots w_n| \leq m$. In particular, L is a finite union of monomials of the form $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$ of degree less than $2m$.*

Lemma 6. *Let $w_1, \ldots, w_n \in \Gamma^*$.*

1. *The monomial $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$ is definable by a $\Sigma_1[<, +1, \min, \max]$-sentence which uses $|w_1 \cdots w_n|$ variables.*
2. *The monomial $w_1 \Gamma^* \cdots w_n \Gamma^*$ is definable by a $\Sigma_1[<, +1, \min]$-sentence which uses $|w_1 \cdots w_n|$ variables.*
3. *The monomial $\Gamma^* w_1 \Gamma^* \cdots w_n \Gamma^*$ is definable by a $\Sigma_1[<, +1]$-sentence which uses $|w_1 \cdots w_n|$ variables.*

Lemma 7. *Let $L \subseteq \Gamma^+$ be a positive Boolean combination of monomials $w_1\Gamma^*w_2 \cdots \Gamma^*w_n$. Then* $\mathrm{Synt}(L) \in [\![x^\omega y x^\omega \le x^\omega]\!]$.

Lemma 8. *Let S be a finite semigroup. For every $w \in S^+$ there exists a factorization $w = x_1 w_1 y_1 \cdots x_m w_m y_m s$ with*

1. $0 \le m \le |S|$ *and* $|x_1 y_1 \cdots x_m y_m s| < 2\,|S|^2 + |S|$,
2. $w_i, s \in S^*$, $x_i, y_i \in S^+$, $|y_i| \le |S|$,
3. $\forall\, i \in \{1,\dots,m\}\ \exists\, e_i \in E(S)$: $x_i = x_i e_i$ *in S and* $y_i = y_i e_i$ *in S.*

Proof. For $w \in S^*$, let $E(w)$ be the set all $e \in E(S)$ such that there exists a factor $x \in S^+$ of w with $|x| \le |S|$ and $xe = x$ in S. We prove the existence of the factorization by induction on $|E(w)|$ with the stronger assertions that $m \le |E(w)|$ and $|x_1 y_1 \cdots x_m y_m s| < 2\,|S|\,|E(w)| + |S|$ instead of condition "1". Suppose $|E(w)| = 0$. By Lemma 1 we have $|w| < |S|$. Hence, we can choose $m = 0$ and $s = w$.

If $|E(w)| \ge 1$, then Lemma 1 yields a nonempty prefix x of w with $|x| \le |S|$ such that $xe = x$ in S for some idempotent $e \in E(S)$. Write $w = xw'$. We have to distinguish two cases. The first case is $e \notin E(w')$. By induction, there exists a factorization $w' = x_1 w_1 y_1 \cdots x_m w_m y_m s$ with $m \le |E(w')| < |E(w)|$ and $|x_1 y_1 \cdots x_m y_m s| \le 2\,|S|\,|E(w')| + |S|$ satisfying conditions "2" and "3". If $m \ge 1$, then $w = (xx_1) w_1 y_1 \cdots x_m w_m y_m s$ is a desired factorization of w. If $w' = s$, then the factorization is $w = xs$ with $m = 0$.

The second case is $e \in E(w')$. Let $w' = w_0 y_0 w''$ such that $y_0 \in S^+$, $|y_0| \le |S|$, $y_0 e = y_0$ in S and $e \notin E(w'')$, i.e., we take y_0 as the last short factor of w' which is stabilized by e. By induction, there exists a factorization $w'' = x_1 w_1 y_1 \cdots x_m w_m y_m s$. Now, $w = x_0 w_0 y_0 \cdots x_m w_m y_m s$ with $x_0 = x$ is a factorization of w of the desired form. ☐

Proposition 9. *Let $L \subseteq \Gamma^+$ be recognized by $S \in [\![x^\omega y x^\omega \le x^\omega]\!]$. Then L is a finite union of monomials $w_1\Gamma^*w_2 \cdots \Gamma^*w_n$ with $n \le |S|^2$ and degree $|w_1 \cdots w_n| < 2\,|S|^3 + |S|^2$.*

Proof. Let $h : \Gamma^+ \to S$ be a homomorphism recognizing L. The order ideal of S generated by a subset $P \subseteq S$ is $\downarrow P = \{x \in S \mid x \le y \text{ for some } y \in P\}$. We define the *depth* of the word $u \in \Gamma^+$ as $d(u) = |\{s \in S \mid h(u) \le_{\mathcal{R}} s\}|$. For every $u \in \Gamma^+$ we are going to construct a language $P_u = w_1\Gamma^*w_2 \cdots \Gamma^*w_n$ with $|w_1 \cdots w_n| < 2d(u)\,|S|^2 + d(u)\,|S|$ such that $u \in P_u \subseteq h^{-1}\big(\downarrow h(u)\big)$. With this claim, $L = \bigcup_{u \in L} P_u$ is a finite union since there are only finitely many monomials of degree less than $2\,|S|^3 + |S|^2$.

In order to avoid unnecessary case distinctions, we set $P_1 = 1$ and $h(1) >_{\mathcal{R}} h(u)$ for all $u \in \Gamma^+$. Let $u = vw$, $v \in \Gamma^*$, $w \in a\Gamma^*$ such that $h(v) >_{\mathcal{R}} h(va)\ \mathcal{R}\ h(u)$. Now, $d(v) < d(u)$ and hence by induction, there exists a monomial P_v with $v \in P_v \subseteq h^{-1}(\downarrow h(v))$ of degree less than $2d(u)\,|S|^2 + d(u)\,|S| - 2\,|S|^2 - |S|$. By Lemma 8 we find a factorization $w = x_1 u_1 y_1 \cdots x_m u_m y_m s$ such that $|x_1 y_1 \cdots x_m y_m s| < 2\,|S|^2 + |S|$ and for all $i \in \{1,\dots,m\}$ there exists an idempotent e_i with $h(x_i)e_i = h(x_i)$ and $h(y_i)e_i = h(y_i)$. Using Lemma 3 we see $h(u) = h(vw) = h(vx_1 \cdots x_m s)$. Now, define the monomial $P_u = P_v\, x_1\Gamma^* y_1 \cdots x_m \Gamma^* y_m s$ of degree less than $2d(u)\,|S|^2 + d(u)\,|S|$. By construction $u \in P_u$. Consider $v'w' \in P_u$ with $v' \in P_v$ and $w' = x_1 w'_1 y_1 \cdots x_m w'_m y_m s$. We have $h(v') \le h(v)$ and since $ese \le e$ for all $s \in S$ and all $e \in E(S)$ we see that $h(x_i) = h(x_i)e_i \ge h(x_i)e_i h(w'_i y_i)e_i = h(x_i w'_i y_i)$. Therefore, $h(x_1 \cdots x_m s) \ge h(w')$ and $h(u) = h(vx_1 \cdots x_m s) \ge h(v'w')$. ☐

4 Existential first-order logic without min or max

At higher levels of the quantifier alternation hierarchy, it is possible to specify the prefix and the suffix of a word by using successor $+1$ as the only predicate (apart from labels $\lambda(x) = a$ for $a \in \Gamma$). At the level Σ_1, the min-predicate is required to determine prefixes, and max is required for suffixes. We have the following inclusions:

$$\Sigma_1[<] \subsetneq \Sigma_1[<,+1] \begin{array}{c} \subsetneq \\ \\ \subsetneq \end{array} \begin{array}{c} \Sigma_1[<,+1,\min] \\ \\ \Sigma_1[<,+1,\max] \end{array} \begin{array}{c} \subsetneq \\ \\ \subsetneq \end{array} \Sigma_1[<,+1,\min,\max]$$

Pin [12] has given an effective characterization for the class of languages definable $\Sigma_1[<]$. For $\Sigma_1[<,+1,\min,\max]$, decidability follows by a result of Pin and Weil [14] (or alternatively by Theorem 4). In this section, we characterize the languages definable in the other fragments and we show that definability within these fragments is decidable. The proofs easily follow from Theorem 4.

Theorem 10. *Let $L \subseteq \Gamma^+$. The following assertions are equivalent:*

1. *L is definable in $\Sigma_1[<,+1,\min]$.*
2. *L is a finite union of monomials $w_1\Gamma^* \cdots w_n\Gamma^*$.*
3. *$\mathrm{Synt}(L) \in [\![x^\omega y x^\omega \leq x^\omega]\!]$ and $h_L(L)$ is a right ideal of $\mathrm{Synt}(L)$.*

Of course, there also is a left-right dual of the above theorem: A language L is definable in $\Sigma_1[<,+1,\max]$ if and only if L is a union of monomials of the form $\Gamma^*w_1 \cdots \Gamma^*w_n$ if and only if $\mathrm{Synt}(L) \in [\![x^\omega y x^\omega \leq x^\omega]\!]$ and $h_L(L)$ is a left ideal of $\mathrm{Synt}(L)$. The following theorem is the analogue of Theorem 10 with neither min nor max predicates.

Theorem 11. *Let $L \subseteq \Gamma^+$. The following assertions are equivalent:*

1. *L is definable in $\Sigma_1[<,+1]$.*
2. *L is a finite union of monomials $\Gamma^*w_1 \cdots \Gamma^*w_n\Gamma^*$.*
3. *$\mathrm{Synt}(L) \in [\![x^\omega y x^\omega \leq x^\omega]\!]$ and $h_L(L)$ is an ideal of $\mathrm{Synt}(L)$.*

Corollary 12. *Let $L \subseteq \Gamma^+$ be a regular language. It is decidable whether L is definable in $\Sigma_1[<,+1]$ (resp. $\Sigma_1[<,+1,\min]$, resp. $\Sigma_1[<,+1,\max]$).*

5 Dot-depth one

A language $L \subseteq \Gamma^+$ has dot-depth one if it is a Boolean combination of monomials of the form $w_1\Gamma^*w_2 \cdots \Gamma^*w_n$ with $w_i \in \Gamma^*$. Knast [6] has shown that a language L has dot-depth one if and only if $\mathrm{Synt}(L) \in \mathbf{B_1}$. Since the latter property is decidable, this gives decidability of dot-depth one. Later, Thérien [22] gave a simpler proof for Knast's result. Both proofs are based on an algebraic concept called *finite categories*, see [24]. In this section, we give a new (more combinatorial) proof of this theorem. The same techniques were used by the authors in order to obtain a characterization for languages of dot-depth one over infinite words [7]. As for dot-depth 1/2, the main advantage of the current proof is that the bounds involved are more explicit.

Theorem 13 (Knast [6], Thomas [23]). *Let $L \subseteq \Gamma^+$. The following assertions are equivalent:*

1. *L is definable in $\mathbb{B}\Sigma_1[<, +1, \min, \max]$.*
2. *L is a Boolean combination of monomials $w_1\Gamma^* w_2 \cdots \Gamma^* w_n$.*
3. *$\mathrm{Synt}(L) \in \mathbf{B_1}$.*

As for dot-depth 1/2, the equivalence of $\mathbb{B}\Sigma_1[<, +1, \min, \max]$ and dot-depth one is due to a result by Thomas [23]. The remainder of this section is devoted to the proof of the above theorem. The following lemma will serve as the link between the algebraic properties of $\mathbf{B_1}$ and the combinatorial properties in Lemma 15 below.

Lemma 14. *Let $S \in \mathbf{B_1}$ and let $k \geq |S| + 1$. For all $a \in \Gamma$ and all $u, v \in S^+$ with $|v| \geq k - 1$ we have $\mathrm{alph}_k(v) \neq \mathrm{alph}_k(va)$ if $u \mathrel{\mathcal{R}} uv >_{\mathcal{R}} uva$.*

Proof. Assume $u \mathrel{\mathcal{R}} uv >_{\mathcal{R}} uva$ and $\mathrm{alph}_k(v) = \mathrm{alph}_k(va)$. Let $va = v'wa$ with $|wa| = k$. Since $wa \in \mathrm{alph}_k(va) = \mathrm{alph}_k(v)$ we have $v = pwaq$ for some $p, q \in S^*$. Let $x = up$, $y = uv'$, and $wa = a_1 \cdots a_k$ for $a_i \in S$. By Lemma 1 there exist $i \in \{1, \ldots, |S|\}$ and $e \in E(S)$ such that $a_1 \cdots a_i = a_1 \cdots a_i e$ in S. In particular $i \leq k - 1$ and $xa_1 \cdots a_i \mathrel{\mathcal{R}} ya_1 \cdots a_i$. Lemma 3 yields $xa_1 \cdots a_i = ya_1 \cdots a_i$ in S. Thus $uva = ywa = xwa \mathrel{\mathcal{R}} u$ in S, a contradiction. $\square$

The following lemma is the main combinatorial ingredient for our proof of Knast's Theorem. It generalizes an idea of Klíma [5] to factors of words. The determinacy mechanism is similar to unambiguous interval logic with lookaround [9].

Lemma 15. *Let $x_i, y_i, u_i, u'_i, v_i, v'_i \in \Gamma^+$ and $u_k, v_k, u'_1, v'_1 \in \Gamma^*$, and let*

$$\begin{aligned} u &= x_1 u_1 \cdots x_k u_k = u'_1 y_1 \cdots u'_\ell y_\ell \\ v &= x_1 v_1 \cdots x_k v_k \ = v'_1 y_1 \cdots v'_\ell y_\ell \end{aligned}$$

such that $x_1 u_1 \cdots x_k$ (resp. $x_1 v_1 \cdots x_k$) is the shortest prefix of u (resp. v) in $x_1\Gamma^+ x_2 \cdots \Gamma^+ x_k$, and $y_1 \cdots u'_\ell y_\ell$ (resp. $y_1 \cdots v'_\ell y_\ell$) is the shortest suffix of u (resp. v) in $y_1\Gamma^+ y_2 \cdots \Gamma^+ y_\ell$.

If u and v are contained in the same languages $w_1\Gamma^+ w_2 \cdots \Gamma^+ w_n$ with $n \leq k + \ell$ and $|w_1 \cdots w_n| \leq |x_1 \cdots x_k\, y_1 \cdots y_\ell|$, then the relative positions of x_k and y_1 are the same in u as in v. More precisely,

1. *$x_1 u_1 \cdots x_k$ is a prefix of $u'_1 \;\Leftrightarrow\; x_1 v_1 \cdots x_k$ is a prefix of v'_1,*
2. *if x_k and y_1 overlap in u or in v, then they have the same overlap in both words,*
3. *$u'_1 y_1$ is a prefix of $x_1 \cdots u_{k-1} \;\Leftrightarrow\; v'_1 y_1$ is a prefix of $x_1 \cdots v_{k-1}$.*

Lemma 16. *Let $S \in \mathbf{B_1}$. For all $u, v, x, s \in S$ and all idempotents $e, f \in S$ we have $u \mathrel{\mathcal{R}} uexf,\ esfv \mathrel{\mathcal{L}} v \;\Rightarrow\; uexfv = uesfv$.*

Proof. Since $u \mathrel{\mathcal{R}} uexf$ and $v \mathrel{\mathcal{L}} esfv$, there exist $y, t \in M$ with $u = uexfy$ and $v = tesfv$. In particular, $u = u(exfy)^\omega$ and $v = (tesf)^\omega v$. We conclude $uexfv = u(exfy)^\omega exf(tesf)^\omega v = u(exfy)^\omega esf(tesf)^\omega v = uesfv$, where the second equality uses $S \in \mathbf{B_1}$. $\square$

Proposition 17. *Let $L \subseteq \Gamma^+$ be recognized by a homomorphism $h : \Gamma^+ \to S$ with $S \in \mathbf{B_1}$ and let $u, v \in \Gamma^+$. If u and v are contained in the same languages $w_1 \Gamma^+ w_2 \cdots \Gamma^+ w_n$ with $n \leq 2\,|S|$ and $|w_1 \cdots w_n| \leq 4\,|S|^2 - 2\,|S|$, then $h(u) = h(v)$.*

Proof. This proof was inspired by Klíma's proof [5] of Simon's Theorem on piecewise-testable languages. The outline is as follows. We consider factorizations induced by the $\mathcal{R}$-factorization of u and the $\mathcal{L}$-factorization of v. Then we transfer the factorization of u to v and vice versa such that the respective orders of the factors in u and v are the same. Finally, we transform v into u by a sequence of h-invariant substitutions.

Consider the $\mathcal{R}$-factorization $u = a_1 u_1 \cdots a_k u_k$ such that

$$h(a_1 u_1 \cdots a_i) \;\mathcal{R}\; h(a_1 u_1 \cdots a_i u_i) >_{\mathcal{R}} h(a_1 u_1 \cdots a_i u_i a_{i+1})$$

for all i. We have $k \leq |S|$. Let j_i be the position of a_i in the above factorization. We color red all positions of u in all the intervals $[\, j_i - |S| \,;\, j_i + |S| - 1\,]$. In particular, the a_i-positions j_i are red. And in general, there is a neighborhood of size $2\,|S|$ around each a_i which contains only red positions. In the worst case, a_1 is the sole exception. Hence, there are at most $2\,|S|^2 - |S|$ red positions in u. Let R_i be the i-th consecutive factor of red positions. Then $u = R_1 u'_1 \cdots R_{k'} u'_{k'}$ for some $u'_i \in \Gamma^+$, $i < k'$, and $u'_{k'} \in \Gamma^*$. Note that $k' \leq k$ because some intervals could overlap. By Lemma 14, the word $R_1 u'_1 \cdots R_i$ is the shortest prefix of u contained in $R_1 \Gamma^+ \cdots R_i$.

Symmetrically, we consider the $\mathcal{L}$-factorization $v = v_1 b_1 \cdots v_\ell b_\ell$ such that

$$h(b_{i-1} v_i b_i \cdots v_\ell b_\ell) <_{\mathcal{L}} h(v_i b_i \cdots v_\ell b_\ell) \;\mathcal{L}\; h(b_i \cdots v_\ell b_\ell)$$

for all i. Let j'_i be the position of b_i in the above factorization. We color blue all positions of v in all the intervals $[\, j'_i - |S| + 1;\, j'_i + |S|\,]$. As before, there are at most $2\,|S|^2 - |S|$ blue positions. Let B_i be the i-th consecutive factor of blue positions. Then $v = v'_1 B_1 \cdots v'_{\ell'} B_{\ell'}$ for $\ell' \leq |S|$ and some $v'_i \in \Gamma^+$, $i > 1$ and $v'_1 \in \Gamma^*$. As before, $B_i \cdots v'_{\ell'} B_{\ell'}$ is the shortest suffix of v contained in $B_i \cdots \Gamma^+ B_{\ell'}$.

Next, we transfer the red positions of u to v, and we transfer the blue positions of v to u. By assumption, $v \in R_1 \Gamma^+ \cdots R_{k'} \Gamma^+$. Therefore, there exists a factorization $v = R_1 v''_1 \cdots R_{k'} v''_{k'}$ such that $R_1 v''_1 \cdots R_i$ is the shortest prefix of v contained in $R_1 \Gamma^+ \cdots R_i$. We color the positions of the R_i's in v red. Similarly, there exists a factorization $u = u''_1 B_1 \cdots u''_{\ell'} B_{\ell'}$ such that $B_i \cdots u''_{\ell'} B_{\ell'}$ is the shortest suffix of u contained in $B_i \cdots \Gamma^+ B_{\ell'}$. We color the positions of the B_i's in u blue. Now, colored positions in u and v are either red or blue or both. By Lemma 15, the colored positions in u have the same order as the colored positions in v. Let w_i be the i-th consecutive factor of colored (red or blue) positions, and write

$$u = w_1 x_1 \cdots w_{n-1} x_{n-1} w_n,$$
$$v = w_1 s_1 \cdots w_{n-1} s_{n-1} w_n.$$

By Lemma 1 and its left-right dual, there exist $e_1, \ldots, e_{n-1} \in E(S)$ and $f_2, \ldots, f_n \in E(S)$ such that each w_i admits a factorization $w_i = p_i r_i q_i$ with $|p_i| \leq |S| - 1$ and $|q_i| \leq |S| - 1$

satisfying

$$\begin{aligned} h(r_i) &= h(r_i)\, e_i && \text{for } 1 \leq i < n, \\ h(r_i) &= f_i\, h(r_i) && \text{for } 1 < i \leq n. \end{aligned}$$

In particular, we can assume $p_1 = 1 = q_n$. Let $x_i' = q_i x_i p_{i+1}$ and $s_i' = q_i s_i p_{i+1}$ for $1 \leq i < n$. Then

$$\begin{aligned} u &= r_1 x_1' r_2 \cdots x_{n-1}' r_n, \\ v &= r_1 s_1' r_2 \cdots s_{n-1}' r_n, \end{aligned}$$

and the r_i's in u cover the positions of the $\mathcal{R}$-factorization of u, whereas the r_i's in v cover the positions of the $\mathcal{L}$-factorization of v. Therefore, for all $1 \leq i < n$

$$\begin{aligned} h(r_1 x_1' \cdots r_i) \;\;&\mathcal{R}\;\; h(r_1 x_1' \cdots r_i) \cdot e_i h(x_i') f_{i+1}, \\ h(r_{i+1} \cdots s_n' r_n) \;\;&\mathcal{L}\;\; e_i h(s_i') f_{i+1} \cdot h(r_{i+1} \cdots s_n' r_n). \end{aligned}$$

By an $(n-1)$-fold application of Lemma 16 we obtain

$$\begin{aligned} h(v) &= h(r_1 s_1' r_2 s_2' r_3 \cdots s_{n-1}' r_n) \\ &= h(r_1 x_1' r_2 s_2' r_3 \cdots s_{n-1}' r_n) \\ &= h(r_1 x_1' r_2 x_2' r_3 \cdots s_{n-1}' r_n) \\ &\;\;\vdots \\ &= h(r_1 x_1' r_2 x_2' r_3 \cdots x_{n-1}' r_n) = h(u) \end{aligned}$$

Note that the substitution rules $s_i' \to x_i'$ are h-invariant in their respective contexts only when applied from left to right. $\square$

Corollary 18. *Let $L \subseteq \Gamma^+$ be recognized by a finite semigroup $S \in \mathbf{B_1}$ and let $u, v \in \Gamma^+$. If u and v are contained in the same monomials $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$ with $n \leq 2\,|S|$ and degree $|w_1 \cdots w_n| < 4\,|S|^2$, then $h(u) = h(v)$.*

We are now ready to prove Theorem 13.

Proof (Theorem 13). "$1 \Leftrightarrow 2$": This follows from Theorem 4.

"$2 \Rightarrow 3$": By Lemma 7 the syntactic semigroup of every monomial $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$ satisfies $x^\omega y x^\omega \leq x^\omega$ and by Lemma 2 it is in $\mathbf{B_1}$. Thus L is recognizable by a direct product $S \in \mathbf{B_1}$ of such semigroups. Since $\mathrm{Synt}(L)$ is a divisor of S, we see that $\mathrm{Synt}(L) \in \mathbf{B_1}$, cf. [11].

"$3 \Rightarrow 2$": Let L be recognized by $h : \Gamma^+ \to S \in \mathbf{B_1}$. We write $u \equiv v$ if u and v are contained in the same monomials of the form $w_1 \Gamma^* w_2 \cdots \Gamma^* w_n$ of degree at most $4\,|S|^2$. We have $L = h^{-1}(P)$ for $P = h(L)$. Corollary 18 shows that every set $h^{-1}(p)$ is a union of $\equiv$-classes. Moreover, $\equiv$ has finite index since there are only finitely many monomials of bounded degree. Every $\equiv$-class is a finite Boolean combination of the required form by specifying which monomials hold and which do not hold. $\square$

6 Dot-depth one without min or max

As for existential first-order logic, one cannot define min- or max-predicates in $\mathbb{B}\Sigma_1[<,+1]$. Therefore, the following inclusions hold:

$$\begin{array}{ccccccc} \mathbb{B}\Sigma_1[<] & \subsetneq & & \subsetneq & \mathbb{B}\Sigma_1[<,+1,\min] & \subsetneq & \\ & & \mathbb{B}\Sigma_1[<,+1] & & & & \mathbb{B}\Sigma_1[<,+1,\min,\max] \\ \mathbb{B}\Sigma_1[+1] & \subsetneq & & \subsetneq & \mathbb{B}\Sigma_1[<,+1,\max] & \subsetneq & \end{array}$$

Simon's Theorem on piecewise testable languages [17] gives decidability of $\mathbb{B}\Sigma_1[<]$. An effective characterization of $\mathbb{B}\Sigma_1[+1]$ is due to Pin [13]. For the fragment $\mathbb{B}\Sigma_1[<,+1,\min,\max]$, decidability follows by Knast's Theorem [6], see Theorem 13. In this section, we give effective characterizations of the remaining fragments. Moreover, we obtain natural subclasses of dot-depth one for the languages definable by the above fragments.

Apart from Theorem 13, the following lemma is the main ingredient in the proof of Theorem 20 below.

Lemma 19. *Let $h : \Gamma^+ \to S \in \mathbf{B_1}$ and let $u, v \in \Gamma^+$. If u and v are contained in the same monomials $w_1\Gamma^* \cdots w_n\Gamma^*$ with $|w_1 \cdots w_n| < 8\,|S|^2$, then $h(u)\ \mathcal{R}\ h(v)$.*

Proof. We write $u \equiv_m v$ if u and v are contained in the same monomials $w_1\Gamma^* w_2 \cdots \Gamma^* w_n$ of degree $|w_1 \cdots w_n| \le m$. Analogously, we write $u \sim_m v$ if u and v are contained in the same monomials $w_1\Gamma^* \cdots w_n\Gamma^*$ of degree $|w_1 \cdots w_n| \le m$. If $u \equiv_m v$ for $m = 4\,|S|^2 - 1$, then by Corollary 18 we have $h(u) = h(v)$.

Let $u \sim_{2m} v$. We want to show $h(u)\ \mathcal{R}\ h(v)$. We can assume $|u|, |v| \ge 2m$, because otherwise $u = v$. Let $u = u'q$ with $|q| = m$. Consider the factorization $v = v'qx$ such that qx is the shortest suffix of v admitting q as a factor, i.e., v is factorized at the last occurrence of q. This factorization exists, since $u \in \Gamma^* q \Gamma^* \ni v$. We claim $u \equiv_m v'q$ and therefore, $h(v) \le_{\mathcal{R}} h(v'q) = h(u)$. Symmetry then yields $h(u)\ \mathcal{R}\ h(v)$.

We now prove the claim. First, let $v'q \in P = w_1\Gamma^* w_2 \cdots \Gamma^* w_n$ with $|w_1 \cdots w_n| \le m$. Then $v \in P\Gamma^*$ and $u \in P\Gamma^*$. Since w_n is a suffix of q, we conclude $u \in P$. Next, suppose $u \in P = w_1\Gamma^* w_2 \cdots \Gamma^* w_n$ with $|w_1 \cdots w_n| \le m$. There exists a monomial $Q = v_1\Gamma^* v_2 \cdots \Gamma^* v_\ell$ with $|v_1 \cdots v_\ell| \le |w_1 \cdots w_n|$ and $u' \in Q \subseteq Pq^{-1}$. Since $u'q \in Qq\Gamma^*$ and the degree of the monomial $Qq\Gamma^*$ is at most $2m$, we obtain $v \in Qq\Gamma^*$. By choice of x we have $v'q \in Qq\Gamma^* \subseteq P\Gamma^*$. Since w_n is a suffix of q, we conclude $v'q \in w_1\Gamma^* w_2 \cdots \Gamma^* w_n$. $\square$

Theorem 20. *Let $L \subseteq \Gamma^+$. The following assertions are equivalent:*

1. *L is definable in $\mathbb{B}\Sigma_1[<,+1,\min]$.*
2. *L is a Boolean combination of monomials $w_1\Gamma^* \cdots w_n\Gamma^*$.*
3. $\mathrm{Synt}(L) \in \mathbf{B_1}$ *and the syntactic homomorphism* $h_L : \Gamma^+ \to \mathrm{Synt}(L)$ *has the property that $h_L(L)$ is a union of $\mathcal{R}$-classes.*

There also is a left-right dual of the above theorem: A language L is definable in $\mathbb{B}\Sigma_1[<,+1,\max]$ if and only if L is a Boolean combination of monomials $\Gamma^* w_1 \cdots \Gamma^* w_n$ if and only if $\mathrm{Synt}(L) \in \mathbf{B_1}$ and $h_L(L)$ is a union of $\mathcal{L}$-classes. Next, we consider the fragment $\mathbb{B}\Sigma_1[<,+1]$ with neither min nor max.

Lemma 21. *Let $h : \Gamma^+ \to S \in \mathbf{B_1}$ and let $u, v \in \Gamma^+$. If u and v are contained in the same monomials $\Gamma^* w_1 \Gamma^* \cdots w_n \Gamma^*$ with $|w_1 \cdots w_n| < 12\,|S|^2$, then $h(u) \mathrel{\mathcal{J}} h(v)$.*

Theorem 22. *Let $L \subseteq \Gamma^+$. The following assertions are equivalent:*

1. *L is definable in $\mathbb{B}\Sigma_1[<, +1]$.*
2. *L is a Boolean combination of monomials $\Gamma^* w_1 \cdots \Gamma^* w_n \Gamma^*$.*
3. *$\mathrm{Synt}(L) \in \mathbf{B_1}$ and the syntactic homomorphism $h_L : \Gamma^+ \to \mathrm{Synt}(L)$ has the property that $h_L(L)$ is a union of $\mathcal{J}$-classes.*

The condition of $h_L(L)$ being a union of $\mathcal{J}$-classes in Theorem 22 has also been used by Beauquier and Pin for an effective characterization of strongly locally testable languages [1].

Corollary 23. *Let $L \subseteq \Gamma^+$ be a regular language. It is decidable whether L is definable in $\mathbb{B}\Sigma_1[<, +1]$ (resp. $\mathbb{B}\Sigma_1[<, +1, \min]$, resp. $\mathbb{B}\Sigma_1[<, +1, \max]$).*

Since in every finite semigroup Green's relation $\mathcal{J}$ is the finest equivalence relation containing both $\mathcal{R}$ and $\mathcal{L}$, we obtain the following corollary.

Corollary 24. *A language $L \subseteq \Gamma^+$ is definable in $\mathbb{B}\Sigma_1[<, +1]$ if and only if L is definable in both $\mathbb{B}\Sigma_1[<, +1, \min]$ and $\mathbb{B}\Sigma_1[<, +1, \max]$.*

7 Summary

We considered subclasses of languages with dot-depth 1/2 and of languages with dot-depth one. These subclasses admit counterparts in terms of fragments of existential first-order logic Σ_1 and its Boolean closure $\mathbb{B}\Sigma_1$. For all fragments, we gave effective algebraic characterizations. We summarize the main results of this paper in Table 1. To shorten notation, we write $\mathbf{B_{1/2}}$ instead of $[\![x^\omega y x^\omega \leq x^\omega]\!]$.

In addition, we gave new proofs for Pin and Weil's Theorem on dot-depth 1/2 and for Knast's Theorem on dot-depth one. The proofs are combinatorial and they improve the bounds involved in computing a language description for a given recognizing semigroup.

References

[1] D. Beauquier and J.-É. Pin. Languages and scanners. *Theor. Comput. Sci.*, 84(1):3–21, 1991.

[2] J. A. Brzozowski and R. Knast. The dot-depth hierarchy of star-free languages is infinite. *J. Comput. Syst. Sci.*, 16(1):37–55, 1978.

[3] R. S. Cohen and J. A. Brzozowski. Dot-depth of star-free events. *J. Comput. Syst. Sci.*, 5(1):1–16, 1971.

[4] Ch. Glaßer and H. Schmitz. Languages of dot-depth 3/2. *Theory Comput. Syst.*, 42(2):256–286, 2008.

[5] O. Klíma. Piecewise testable languages via combinatorics on words. Preprint, 2009.

[6] R. Knast. A semigroup characterization of dot-depth one languages. *RAIRO, Inf. Théor.*, 17(4):321–330, 1983.

Languages	**Logics**	**Algebra**	
$\bigcup w_1\Gamma^* w_2 \cdots \Gamma^* w_n$	$\Sigma_1[<,+1,\min,\max]$	$\mathbf{B_{1/2}}$	[14], Thm. 4
$\bigcup w_1\Gamma^* \cdots w_n\Gamma^*$	$\Sigma_1[<,+1,\min]$	right ideals in $\mathbf{B_{1/2}}$	Thm. 10
$\bigcup \Gamma^* w_1 \cdots \Gamma^* w_n$	$\Sigma_1[<,+1,\max]$	left ideals in $\mathbf{B_{1/2}}$	cf. Thm. 10
$\bigcup \Gamma^* w_1 \cdots \Gamma^* w_n\Gamma^*$	$\Sigma_1[<,+1]$	ideals in $\mathbf{B_{1/2}}$	Thm. 11
$\mathbb{B}(w_1\Gamma^* w_2 \cdots \Gamma^* w_n)$	$\mathbb{B}\Sigma_1[<,+1,\min,\max]$	$\mathbf{B_1}$	[6], Thm. 13
$\mathbb{B}(w_1\Gamma^* \cdots w_n\Gamma^*)$	$\mathbb{B}\Sigma_1[<,+1,\min]$	$\mathcal{R}$-classes in $\mathbf{B_1}$	Thm. 20
$\mathbb{B}(\Gamma^* w_1 \cdots \Gamma^* w_n)$	$\mathbb{B}\Sigma_1[<,+1,\max]$	$\mathcal{L}$-classes in $\mathbf{B_1}$	cf. Thm. 20
$\mathbb{B}(\Gamma^* w_1 \cdots \Gamma^* w_n\Gamma^*)$	$\mathbb{B}\Sigma_1[<,+1]$	$\mathcal{J}$-classes in $\mathbf{B_1}$	Thm. 22

Table 1: Languages around dot-depth one.

[7] M. Kufleitner and A. Lauser. Languages of dot-depth one over infinite words. In *LICS'11, Proceedings*, pages 23–32, IEEE, 2011.

[8] M. Kufleitner and A. Lauser. Around dot-depth one. Technical report no. 2011/03, Universität Stuttgart, Informatik, 2011.

[9] K. Lodaya, P. K. Pandya, and S. S. Shah. Around dot depth two. In *DLT'10, Proceedings*, volume 6224 of *LNCS*, pages 305–316. Springer, 2010.

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

[11] J.-É. Pin. *Varieties of Formal Languages*. North Oxford Academic, London, 1986.

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

[13] J.-É. Pin. Expressive power of existential first-order sentences of Büchi's sequential calculus. *Discrete Math.*, 291(1-3):155–174, 2005.

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

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

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

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

[18] I. Simon. Factorization forests of finite height. *Theor. Comput. Sci.*, 72(1):65–94, 1990.

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

[20] H. Straubing. Finite semigroup varieties of the form $\mathbf{V}*\mathbf{D}$. *J. Pure Appl. Algebra*, 36(1):53–94, 1985.

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

[22] D. Thérien. Categories et langages de dot-depth un. *RAIRO, Inf. Théor.*, 22(4):437–445, 1988.

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

[24] B. Tilson. Categories as algebras: An essential ingredient in the theory of monoids. *J. Pure Appl. Algebra*, 48:83–198, 1987.