Games, equations and dot-depth two monoids

By: F. Blanchet-Sadri

Blanchet-Sadri, F. (1992) Games, equations and dot-depth two monoids, *Discrete Applied Mathematics* 39, 99-111.

Abstract:
Given any finite alphabet A and positive integers $m_1, \ldots, m_k$, congruences on A^*, denoted by $\sim(m_1, \ldots, m_k)$ and related to a version of the Ehrenfeucht-Fraisse game, are defined. Level k of the Straubing hierarchy of aperiodic monoids can be characterized in terms of the monoids $A^*/{\sim}(m_1, \ldots m_k)$. A natural subhierarchy of level 2 and equation systems satisfied in the corresponding varieties of monoids are defined. For $A \geq 2$, a necessary and sufficient condition is given for $A^*/{\sim}(m_1, \ldots, m_k)$ to be of dot-depth exactly 2. Upper and lower bounds on the dot-depth of the $A^*/{\sim}(m_1, \ldots m_k)$ are discussed.

Article:

1. Introduction

In this paper, we present results relative to the characterization of dot-depth k monoids. This topic is of interest from the points of view of formal language theory, symbolic logic and complexity of boolean circuits. The results are obtained by a technical and detailed use of a version of the Ehrenfeucht-Fraisse game.

Let A be a given finite alphabet. The regular languages over A are those subsets of A^*, the free monoid generated by A, constructed from the finite languages over A by the boolean operations, the concatenation product and the star. The star-free languages are those regular languages which can be obtained from the finite languages by the boolean operations and the concatenation product only. According to Schützenberger [17], $L \subseteq A^*$ is star-free if and only if its syntactic monoid $M(L)$ is finite and aperiodic. General references on the star-free languages are McNaughton and Papert [12], Eilenberg [8] or Pin [14].

Natural classifications of the star-free languages are obtained based on the alternating use of the boolean operations and the concatenation product. Let $A^+ = A^* \setminus \{1\}$, where 1 denotes the empty word. Let

$$A^+\mathscr{B}_0 = \{L \subseteq A^+ \mid L \text{ is finite or cofinite}\},$$

$$A^+\mathscr{B}_{k+1} = \{L \subseteq A^+ \mid L \text{ is a boolean combination of languages of the form } L_1 \ldots L_n\ (n \geq 1) \text{ with } L_1, \ldots, L_n \in A^+\mathscr{B}_k\}.$$

Only nonempty words over A are considered to define this hierarchy; in particular, the complement operation is applied with respect to A^+. The language classes $A^+\mathcal{B}_0, A^+\mathcal{B}_1$ ... form the so-called dot-depth hierarchy introduced by Cohen and Brzozowski in [6]. The union of the classes $A^+\mathcal{B}_0, A^+\mathcal{B}_1$, ... is the class of star-free languages.

Our attention is directed toward a closely related and more fundamental hierarchy, this one in A^*, introduced by Straubing in [20]. Let

$$A^*\mathscr{V}_0 = \{\emptyset, A^*\} \quad \text{where } \emptyset \text{ is the empty set},$$

$$A^*\mathscr{V}_{k+1} = \{L \subseteq A^* \mid L \text{ is a boolean combination of languages of the form } L_0 a_1 L_1 a_2 \ldots a_n L_n\ (n \geq 0) \text{ with } L_0, \ldots, L_n \in A^*\mathscr{V}_k \text{ and } a_1, \ldots, a_n \in A\}.$$

Let $A^*\mathcal{V} = \bigcup_{k \geq 0} A^*\mathcal{V}_k$. $L \subseteq A^*$ is star-free if and only if $L \in A^*\mathcal{V}_k$ for some $k \geq 0$. The dot-depth of L is the smallest such k.

For k ≥ 1, let us define subhierarchies of $A^*\mathcal{V}$ as follows: for all $m \geq 1$, let

$$A^*\mathcal{V}_{k,m} = \{L \subseteq A^* \mid L \text{ is a boolean combination of languages of the form } L_0a_1L_1a_2\ldots a_nL_n\ (0 \leq n \leq m) \text{ with } L_0,\ldots,L_n \in A^*\mathcal{V}_{k-1} \text{ and } a_1,\ldots,a_n \in A\}.$$

We have $A^*\mathcal{V}_k = \bigcup_{m\geq 1} A^*\mathcal{V}_{k,m}$. Easily, $A^*\mathcal{V}_{k,m} \subseteq A^*\mathcal{V}_{k+1,m}$, $A^*\mathcal{V}_{k,m} \subseteq A^*\mathcal{V}_{k,m+1}$. Similarly, subhierarchies of $A^+\mathcal{B}_k$ can be defined. In $A^+\mathcal{B}_1$ several hierarchies and classes of languages have been studied; the most prominent examples are the β-hierarchy [5], also called depth-one finite cofinite hierarchy, and the class of locally testable languages.

$\mathcal{W}$ is a *-variety of languages if

(1) for every finite alphabet A, $A^*\mathcal{W}$ denotes a class of recognizable (recognizable means recognizable by a finite automaton or regular) languages over A closed under boolean operations,

(2) if $L \in A^*\mathcal{W}$ and $a \in A$, then $a^{-1}L = \{w \in A^* \mid aw \in L\}$ and $La^{-1} = \{w \in A^* \mid wa \in L\}$ are in $A^*\mathcal{W}$, and

(3) if $L \in A^*\mathcal{W}$ and $\varphi: B^* \to A^*$ is a morphism, then $L\varphi^{-1} = \{ w \in B^* \mid w\varphi \in L\} \in B^*\mathcal{W}$.

Eilenberg [8] has shown that there exists a one-to-one correspondence between *-varieties of languages and some classes of finite monoids called *M*-varieties. *W* is an *M*-variety if

(1) it is a class of finite monoids closed under division, i.e., if $M \in W$ and $M' < M$ (< denotes the divide relationship between monoids), then $M' \in W$, and

(2) it is closed under finite direct product, i.e., if $M, M' \in W$, then $M \times M' \in W$.

To a given *-variety of languages $\mathcal{W}$ corresponds the *M*-variety $W = \{M(L) \mid L \in A^*\mathcal{W}$ for some $A\}$ and to a given *M*-variety *W* corresponds the *-variety of languages $\mathcal{W}$ where $A^*\mathcal{W} = \{L \subseteq A^* \mid$ there is $M \in W$ recognizing $L\}$. The Straubing hierarchy gives examples of *-varieties of languages. One can show that $\mathcal{V}$, $\mathcal{V}_k$ and $\mathcal{V}_{k,m}$ are *-varieties of languages. Let the corresponding *M*-varieties be denoted by *V*, V_k and $V_{k,m}$ respectively. *V* is the *M*-variety of aperiodic monoids. We have that for $L \in A^*$, $L \in A^*\mathcal{V}$ if and only if $M(L) \in V$, for each $k \geq 0$, $L \in A^*\mathcal{V}_k$ if and only if $M(L) \in V_k$, and for $k \geq 1$, $m \geq 1$, $L \in A^*\mathcal{V}_{k,m}$ if and only if $M(L) \in V_{k,m}$.

An outstanding open problem is whether one can decide if a star-free language has dot-depth k, i.e., can we effectively characterize the *M*-varieties V_k? The variety V_0 consists of the trivial monoid alone, V_1 of all finite $\mathcal{J}$-trivial monoids [181. Straubing [21] conjectured an effective characterization, based on the syntactic monoid of the language, for the case $k = 2$. His characterization, formulated in terms of a novel use of categories in semigroup theory recently developed by Tilson [24], is shown to be necessary in general, and sufficient for an alphabet of two letters.

In the framework of semigroup theory, Brzozowski and Knast [4] showed that the dot-depth hierarchy is infinite. Thomas [231 gave a new proof of this result, which shows also that the Straubing hierarchy is infinite, based on a logical characterization of the dot-depth hierarchy that he obtained in [221 (Perrin and Pin [131 gave one for the Straubing hierarchy) and the following version of the Ehrenfeucht-Fraisse game.

First, one regards a word $w \in A^*$ of length |w| as a word model $w = \langle\{1, ..., |w|\}, <^w (Q_a^w)_{a\in A}\rangle$ where the universe {1, ..., |w|} represents the set of positions of letters in *w*, $<^w$ denotes the <-relation in *w*, Q_a^w are unary relations over {1, …, |w|} containing the positions with letter *a*, for each $a \in A$. For a sequence $\overline{m} = (m_1, ..., m_k)$ of positive integers, where $k \geq 0$, the game $\mathcal{G}\overline{m}(u, v)$ is played between two players I and II on the word models *u* and *v*. A play of the game consists of *k* moves. In the *i*th move, player I chooses, in *u* or in *v*, a sequence of m_1 positions; then player II chooses, in the remaining word, also a sequence of m_1 positions. After *k* moves, by concatenating the sequences chosen from u and v, two sequences $p_1 \ldots p_n$ from u and $q_1 \ldots q_n$ from v have been formed where $n = m_1 + \ldots + m_k$. Player II has won the play if $p_i <^u p_j$ if and only if $q_i <^v q_j$, and $Q_a^u p_i$ if and only if $Q_a^v q_i$, $a \in A$ for $1 \leq i, j \leq n$. If there is a winning strategy for player II in the game $\mathcal{G}\overline{m}(u, v)$ to win each play we write $u \sim\overline{m}\ v$. $\sim\overline{m}$ naturally defines a congruence on A^* which we denote also by $\sim\overline{m}$. The standard Ehrenfeucht-Fraisse game [7] is the special case $\mathcal{G}(1, \ldots, 1)(u, v)$. Thomas [22,231 and Perrin and Pin [13] infer the following congruence characterization of the A^*V_k and the $A^*V_{k,m}$, i.e., $L \in A^*\mathcal{V}_k$ if and only if L is a union of classes of some $\sim(m_1,...,m_k)$ and $L \in A^*\mathcal{V}_{k,m}$ if and only if L is a union of classes of some $\sim(m, m_2, ..., m_k)$. This implies the following congruence characterization of the V_k and the $V_{k,m}$, i.e., $V_k = \{A^*/\sim | \sim \supseteq \sim(m_1, \ldots, m_k)$ for some m_i, $i = 1,...,k\}$, and $V_{k,m} = \{A^*/\sim | \sim \supseteq \sim(m, m_2, \ldots, m_k)$ for some m_i, $i = 2, \ldots, k\}$. In [2], it was shown that for fixed $(m_1, \ldots, m_k)$, it is decidable if a language is a union of some classes of $\sim(m_1, ..., m_k)$, or, equivalently, it is decidable if the syntactic monoid of a language divides $A^*/\sim(m_1,...,m_k)$.

Let $u, v \in A^*$. A monoid M satisfies the equation $u = v$ if and only if $u\varphi = v\varphi$ for all morphisms $: A^* \to M$. One can show that the class of monoids M satisfying the equation $u = v$ is an M-variety, denoted by $W(u, v)$. Let $(u_n, v_n)_{n>0}$ be a sequence of pairs of words of A^*. Consider the following M-varieties: $W' = \bigcap_{n>0} W(u_n, v_n)$ and $W'' = \bigcup_{m>0} \bigcap_{n \geq m} W(u_n, v_n)$. We say that W' (W'') is defined (ultimately defined) by the equations $u_n = v_n$ ($n>0$): this corresponds to the fact that a monoid M is in W' (W'') if and only if M satisfies the equations $u_n = v_n$ for all $n > 0$ (for all n sufficiently large). The equational approach to varieties is discussed in Eilenberg [8]. Eilenberg showed that every M-variety is ultimately defined by a sequence of equations. For example, the M-variety V of aperiodic monoids is ultimately defined by the equations $x^n = x^{n+1}$ ($n > 0$). The M-variety V_1 is ultimately defined by the equations $(xy)^m = (yx)^m$ and $x^m = x^{m+1}$ ($m > 0$). This gives a decision procedure for V_1, i.e., $M \in V_1$ if and only if for all $x, y \in M$, $(xy)^m = (yx)^m$ and $x^m = x^{m+1}$ with m the cardinality of M. One can show that every M-variety generated by a single monoid is defined by a (finite or infinite) sequence of equations. $V_{1,m}$ being generated by $A^*/\sim(m)$, are the M-varieties $V_{1,m}$ defined by a finite sequence of equations? An attempt to answer this open problem was made in [3]. There, systems of equations were defined which are satisfied in the $V_{t,m}$ ([10,11] provide an equation system for level 1 of the dot-depth hierarchy). It was shown that those equation systems characterize completely $V_{1,1}$, $V_{1,2}$ and $V_{1,3}$. More precisely, $V_{1,1}$ is defined by $x = x^2$ and $xy = yx$, $V_{1,2}$ by $xyzx = xyxzx$ and $(xy)^2 = (yx)^2$, and $V_{1,3}$ by $xzyxvxwy = xzxyxvxwy$, $ywxvxyzx = ywxvxyxzx$ and $(xy)^3 = (yx)^3$.

This paper is concerned with applications of the above mentioned congruence characterization of the V_k and the $V_{k,m}$. Other applications appear in [1-3]. [2] answers a conjecture of Pin [15] concerning tree hierarchies of monoids. The problem of finding equations satisfied in the $V_{2,m}$ problem related to the effective characterization of the $V_{2,m}$ and hence of V_2, is the subject of Section 3. More precisely, systems of equations are defined which are satisfied in the $V_{2,m}$. In Section 4, we are interested in the following question: for an alphabet of at least two letters, find a necessary and sufficient condition for $A^*/\sim(m_1,...,m_k)$ to be of dot-depth exactly d. Such a condition is given for $d = 1$ and $d = 2$. It is also shown that for all sufficiently large m_i, $A^*/\sim(m_1,...,m_k)$ is of dot-depth exactly k. The proofs rely on some properties of the congruences $\sim\overline{m}$ stated in the next section. The reader is referred to the books by Pin [14] and Enderton [9] for all the algebraic and logical terms not defined in this paper.

2. Some properties of the $\sim\overline{m}$

2.1. An induction lemma

The following lemma is a basic result (similar to one in [16] regarding ~(1.....I)) which allows to resolve games with $k + 1$ moves into games with k moves and thereby allows to perform induction arguments. In what follows, $u[1, p)$($u(p, |u|]$) denotes the segment of u to the left (right) of position p and $u(p, q)$ the segment of u between positions p and q.

Lemma 2.1. *Let $\overline{m} = (m_1,...,m_k)$. $u \sim(m, m_1, \ldots, m_k)\ v$ if and only if*

(1) for every $p_1, \ldots, p_m \in u$ ($p_1 \leq \ldots \leq p_m$) there are $q_i,...,q_m \in v$ ($q_1 \leq \ldots \leq q_m$) such that

(i) $Q_a^u p_i$ if and only if $Q_a^v q_i$, $a \in A$ for $1 \leq i \leq m$,

(ii) $u[1, p_1) \sim\overline{m}\ v[1, q_1)$,

(iii) $u(p_i, p_{i+1}) \sim\overline{m}\ v(q_i, q_{i+1})$ *for* $1 \le 1 \le m - 1$,

(iv) $u(p_m, |u|] \sim\overline{m}\ v(qm, |v|]$ *and*

(2) *for every* $q_1, \ldots, q_m \in v(q_1 \le \ldots \le q_m)$ *there are* $p1, \ldots, pm \in u$ $(p_1 \le \ldots \le p_m)$ *such that (i), (ii), (iii) and (iv) hold.*

2.2. An inclusion lemma

Define $\mathcal{N}(m_1, \ldots, m_k) = (m_1+1) \ldots (m_k + 1)\text{-}1$. We can show that $x^N \sim(m_1, \ldots, m_k)\ x^{N+1}$ $(N = \mathcal{N}(m_1, \ldots, m_k))$ and that N is the smallest n such that $x^n \sim(m_1, \ldots, m_k)\ x^{n+1}$ for $|x| = 1$ (the proof is similar to the one of a property of $\sim(1, ..., 1)$ in [23]). It follows that if $u, v \in A^*$ and $u \sim(m_1, \ldots, m_k)\ v$, then $|u|_a = |v|_a < \mathcal{N}(m_1, \ldots, m_k)$ or $|u|_a, |v|_a \ge \mathcal{N}(m_1, \ldots, m_k)$ (here $|w|_a$ denotes the number of occurrences of the letter a in w). The following lemma follows easily from Lemma 2.1 and the above remarks.

Lemma 2.2. $\sim(m_1, \ldots, m_k) \subseteq \sim(\mathcal{N}(m_1, \ldots, m_k))$, *and* $\sim(m_1, \ldots, m_k) \nsubseteq \sim(\mathcal{N}(m_1, \ldots, m_k) + 1)$. *If* $k \le k'$ *and* $\exists 0=j_0 < \ldots < j_{k-1} < j_k = k'$ *such that* $m_i \le \mathcal{N}(m'_j \quad , \ldots, m'_i)$ *for* $1 \le i \le k$, *then* $\sim(m'_1,\ldots, m'_k) \subseteq \sim(m_1, \ldots, m_k)$.

2.3. Some combinatorial lemmas

We will need the following combinatorial properties of the congruences $\sim(m)$.

Lemma 2.3 (Simon [181). *Let* $m \ge 1$. *Let* $u,v \in A^*$. *If* $u \sim(m)v$, *then there exists* $w \in A^*$ *such that u is a subword of w, v is a subword of w and* $u \sim(m)\ w \sim(m)\ v$ *(a word* $a_1 \ldots a_n$ *(where* $a_1,\ldots,a_n$ *are letters) is a subword of w if there exist words* $w_0, \ldots, w_n$ *such that* $w = w_0a_1w_1a_2 \ldots a_nw_n$*).*

Lemma 2.4 (Simon [18]). *Let* $m \ge 1$. *Let* $u,v \in A^*$. *Then*

(1) $u \sim(m)uv$ *if and only if there exist* $u_1,\ldots,u_m \in A^*$ *such that* $u = u_1 \ldots u_m$ *and* $v\alpha \subseteq u_m\alpha \subseteq \ldots \subseteq u_1\alpha$ *(here* $w\alpha$ *denotes the set of letters in w).*

(2) $u \sim(m)\ vu$ *if and only if there exist* $u_1,\ldots,u_m \in A^*$ *such that* $u = u_1 \ldots u_m$ *and* $v\alpha \subseteq u_1\alpha \subseteq \ldots \subseteq u_m\alpha$.

3. Equations and the $V_{2,m}$

Simon's effective characterization of V_1 [18] depends on a detailed study of combinatorial properties of the congruences $\sim(m)$ (like those in Lemmas 2.3 and 2.4). A monoid M in V_1 satisfies $(xy)^m$ —$(yx)^m$ and $x^m = x^{m+1}$ for some m since $M < A^*/\sim(m)$ for some m and $(xy)^m \sim(m)(yx)^m$ and $x^m \sim(m)x^{m+1}$. It turns out that these two equations form a complete system of equations for V_1. Subsection 3.1 studies some combinatorial properties of the congruences $\sim(1, m)$ and gives equations satisfied in $V_{2,1}$.

3.1. Equations and $V_{2,1}$

In the following, we talk about positions spelling the first and last occurrences of every subword of length $\le m$ of a word w. We illustrate what we mean by this with the following example. Let $A = \{a, b, c\}$ and

$w = abccccaabbabbacccabababccaaaabbaa \ldots$.

↑↑↑↑ ↑ ↑ ↑↑↑↑ ↑↑↑↑↑

p

The six arrows on the left point to the positions which spell the first occurrences of every subword of length ≤2 in $w[1,p)$ and the eight arrows on the right (before the one pointing to p) to the positions which spell the last occurrences of every subword of length ≤2 in $w[1,p)$.

Lemma 3.1. *Let* $m \ge 1$. *Let* $u, v \in A^+$ *and let* $p_1, \ldots, p_s$ *in u* $(p_1 < \ldots < p_s)(q_1, \ldots, qs$, *in v* $(q_1 < \ldots < q_s))$ *be the positions which spell the first and last occurrences of every subword of length ≤m in u (v).* $u \sim(1,m)v$ *if and only if*

(1) $s=s'$

(2) $Q^u_a p_i$ *if and only if* $Q^v_a q_i$, $a \in A$ *for* $1 \le i \le s$ *and*

(3) $u(p_i, p_{i+1}) \sim(1)\ v(q_i, q_{i+1})$ *for* $1 \le i \le s - 1$.

Proof. Assume (1), (2) and (3) hold. A winning strategy for player II in the game $\mathcal{G}(1, m)(u, v)$ to win each play is described as follows. Let p be a position in u chosen by player I in the first move (the proof is similar when starting with a position in v). Assume $Q_a^u p$.

Case 1: *p is among $p_1, \ldots, p_s$, i.e., $p=p_i$ for some i, $1 \le i \le s$.* Since (1) holds, we can consider $q = q_i$. (2) implies that $Q_a^v q$.

Case 2: *$p \in u(p_i, p_{i+1})$ for some i, $1 \le Ii \le s - 1$.* From (3), there is $q \in v(q_i, q_{i+1})$ such that $Q_a^v q$. In either case, (1), (2), (3) and the choice of q imply that $u(p, |u| \sim(m)\ v(q, |v|]$ and $u[1, p) \sim(m)\ v[1, q)$.

Conversely, assume $u \sim(1, m)\ v$. (1) and (2) obviously hold. Also, $u(p_i, p_{i+1}) \sim(1)\ v(q_i, q_{i+1})$ for $1 \le i \le s$-1. To see this, let p be in $u(p_i, p_{i+1})$ (the proof is similar when starting with q in $v(q_i, q_{i+1})$). Consider the following play of the game $\mathcal{G}(1, m)(u, v)$. Player I, in the first move, chooses p. Hence there exists q in v such that $u(p, |u|] \sim(m)\ v(q, |v|]$ and $u[1, p) \sim(m)\ v[1, q)$. Assume that $q \notin v(q_i, q_{i+1})$. Hence $q \in v[1, q_i]$ or $q \in v[q_{i+1}, |v|]$. From the choice of the p_i and the q_i, either $u(p, |u|] \nsim(m)\ v(q, |v|]$ or $u[1, p) \nsim(m)\ v[1, q)$. Contradiction. The result follows.

Proposition 3.2. *Let $m \ge 1$. Let $u, v \in A^*$. If $u \sim(1, m)\ u$, then there exists $w \in A^*$ such that u is a subword of w, v is a subword of w and $u \sim(1, m)\ w \sim(1,m)\ v$.*

Proof. Let $A = \{a_1, \ldots, a_r\}$. If $r =1$, $u=v$ or $|u|, |v| \ge \mathcal{N}(1, m)$ by Section 2. For $r > 1$, let $p_1, \ldots, p_s$ ($p_1 < \cdots < p_s$) be the positions which spell the first and last occurrences of every subword of length $\le m$ in u. s is no more than $2m\ (r + 1)^m$. Assume $Q^u_{a_{j_1}} p_i$. Since $u \sim(1,m)\ v$, by Lemma 3.1, the positions $q_1, \ldots, q_s$ ($q_1 < \cdots < q_s$) in v which spell the first and last occurrences of every subword of length $\le m$ in v are such that $Q^v_{a_{j_1}} q_i$ for $1 \le i \le s$ and $u(p_i, p_{i+1}) \sim(1)\ v(q_i, q_{i+1})$ for $1 \le i \le s$-1. Hence by Lemma 2.3, since $u(p_i, p_{i+1}) \sim(1)\ v(q_i, q_{i+1})$, there exists w_i such that $u(p_i, p_{i+1})$ is a subword of w_i, $v(q_i, q_{i+1})$ is a subword of w_i and $u(p_i, p_{i+1}) \sim(1)\ w_i \sim(1)\ v(q_i, q_{i+1})$. Let $w = a_{j_1} w_1 a_{j_2} w_2 \ldots a_{j_{s-1}} w_{s-1} a_{j_s}$. u is a subword of w, v is a subword of w and $u \sim(1, m)\ w \sim(1, m)\ v$ by Lemma 3.1.

Now, let us define classes of equations as follows. For $m \ge 1$, $\mathcal{C}^1_{(1,m)}$ is a class of equations consisting of

$$u_1 \ldots u_m xy v_1 \ldots v_m = u_1 \ldots u_m yx v_1 \ldots v_m$$

where the u and the v are of the form $x^e y$, $y^e x$, xy^e or yx^e for some e, $1 \le e \le \mathcal{N}(1, m)$. The equation $(xy)^m xy(xy)^m = (xy)^m yx(xy)^m$ is an example.

$\mathcal{C}^2_{(1,m)}$ consists of the equations

$$u_1 \ldots u_i x^{m-i} x x^{m-j} v_1 \ldots v_j = u_1 \ldots u_i x^{m-i} x^2 x^{m-j} v_1 \ldots v_j$$

where the u and the v are as above and $0 \le i, j \le m$. The equation $(xy)^m x(xy)^m = (xy)^m x^2 (xy)^m$ is an example.

Note that the equations in $\mathcal{C}^1_{(1,m)}$ are of the form $w_1 xy w_2 = w_1 yx w_2$ and the ones in $\mathcal{C}^2_{(1,m)}$ of the form $w_3 x w_4 = w_3 x^2 w_4$. Recall from Section 1 that $xy = yx$ and $x = x^2$ are the defining equations for $V_{1,1}$.

Theorem 3.3. *Every monoid in $V_{2,1}$ satisfies $\mathcal{C}^1_{(1,m)} \cup \mathcal{C}^2_{(1,m)}$ for all sufficiently large m.*

Proof. It is easily seen, using Lemma 3.1, that monoids in $V_{2,1}$ satisfy $\mathcal{C}^1_{(1,m)} \cup \mathcal{C}^2_{(1,m)}$ for some $m \ge 1$. This comes from the fact that if $M \in V_{2,1}$, then $M < A^* /\sim(1, m)$ for some $m \ge 1$. Since $A^*/\sim(1,m)$ satisfies $\mathcal{C}^1_{(1,m)} \cup \mathcal{C}^2_{(1,m)}$, M satisfies $\mathcal{C}^1_{(1,m)} \cup \mathcal{C}^2_{(1,m)}$. Moreover, if M in $V_{2,1}$ satisfies $\mathcal{C}^1_{(1,m)} \cup \mathcal{C}^2_{(1,m)}$ for some $m \ge 1$, then it satisfies $\mathcal{C}^1_{(1,n)} \cup \mathcal{C}^2_{(1,n)}$ for all $n \ge m$ since $\sim(1, n) \subseteq \sim(1, m)$ for those n.

3.2. Equations and the $V_{2,m}$ where $m > 1$

This subsection generalizes the equation systems of the preceding subsection so that the generalized equations are satisfied in the $V_{2,m}$.

Lemma 3.4. *Let* $m_1>1$, $m_2 \geq 1$. *Let* $u,v \in A^+$ *and let* $p_1, \ldots, p_s$ *in* u $(p_1 < \cdots < p_s)$ $q_1,\ldots,q_{s'}$, *in* v $(q_1 < \cdots < q_{s'})$ *be the positions which spell the first and last occurrences of every subword of length* $\leq m_2$ *in* u (v). $u \sim(m_1,m_2)v$ *if and only if*

(1) $s=s'$,

(2) $Q_a^u p_i$ *if and only if* $Q_a^v q_i$, $a \in A$ *for* $1 \leq i \leq s$,

(3) $u(p_i,p_{i+1}) \sim(m_1\text{-}2,m_2)v(q_i, q_{i+1})$ *for* $1 \leq i \leq s - 1$,

(4) for $1 \leq i \leq s\text{-}1$ *and for every* $p'_1, \ldots, p'_{m_1-1} \in u(p_i, p_{i+1})$ $(p'_1 < \ldots < p'_{m_1-1})$, *there exist* $q'_1, \ldots, q'_{m_1-1} \in v(q_i,q_{i+1})$ $(q'_1 < \ldots < q'_{m_1-1})$ *such that*

(1') $Q_a^u p'_j$ *if and only if* $Q_a^v q'_j$, $a \in A$ *for* $1 \leq j \leq m_1 - 1$,

(2') $u(p'_j, p'_{j+1}) \sim(m_2)\ v(q'_j, q'_{j+1})$ *for* $1 \leq j \leq m_1\text{-}\ 2$ *and*

(3') $u(pi, p'_1) \sim(m_2)\ v(qi, q'_1)$.

Also, there exist $q'_1,\ldots,q'_{m_1-1} \in v(q_i, q_{i+1})$ *(which may be different from the positions which satisfy (1'), (2') and (3'))* $(q'_1 < \ldots < q'_{m_1-1})$ *such that (1'), (2') and (3'')* $u(p'_{m_1-1}, p_{i+1}) \sim(m_2)\ v(q'_{m_1-1}, q_{i+1})$ *hold. Similarly, for every* $q'_1, \ldots, q'_{m_1-1} \in v(q_i, q_{i+1})$ $(q'_1 < \ldots < q'_{m_1-1})$, *there exist* $p'_1, \ldots, p'_{m_1-1} \in u(p_i, p_{i+1})$ $(p'_1 < \ldots < p'_{m_1-1})$ *such that (1'), (2'), (3') hold (also (1'), (2'), (3'') hold) and*

(5) for $1 \leq i \leq s - 1$ *and for every* $p'_1,\ldots, p'_{m_1} \in u(p_i, p_{i+1})$ $(p'_1 < \ldots < p'_{m_1})$, *there exist* $q'_1,\ldots, q'_{m_1} \in v(q_i, q_{i+1})$ $(q'_1 < \ldots < q'_{m_1})$ *such that*

(1´´´) $Q_a^u p'_j$ *if and only if* $Q_a^v q'_j$, $a \in A$ *for* $1 \leq j \leq m_1$ *and*

(2´´´) $u(p'_j, p'_{j+1}) \sim(m_2)\ v(q'_j, q'_{j+1})$ *for* $1 \leq j \leq m_1 - 1$.

Similarly, for every $q'_1,\ldots, q'_{m_1} \in v(q_i, q_{i+1})$ $(q'_1 < \ldots < q'_{m_1})$, *there exist* $p'_1, \ldots, p'_{m_1} \in u(p_i, p_{i+1})(p'_1 < \ldots < p'_{m_1})$ *such that (1''') and (2''') hold.*

Proof. Assume (1), (2), (3), (4) and (5) hold. A winning strategy for player II in the game $\mathcal{G}(m_1,m_2)(u, v)$ to win each play is described as follows. Let $p'_1, \ldots, p'_{m_1}(p'_1 \leq \cdots \leq p'_{m_1})$ be positions in u chosen by player I in the first move (the proof is similar when starting with positions in v).

Case 1. If some of the p'_j are among $p_1, \ldots, p_s$, then for such a p'_j, there exists i_j, $1 \leq i_j \leq s$ such that $p'_j = p_{i_j}$. For such a p'_j, since (1) holds, we may consider $q'_j = q_{i_j}$. (2) implies that $Q_a^u p'_j$ if and only if $Q_a^v q'_j$.

Case 2. If $p'_j, p'_{j+1}, \ldots, p'_{j+l} \in u(p_i, p_{i+1})$ for some i, $1 \leq i \leq s\text{-}\ 1$, $1 \leq j \leq j + 1 \leq m_l$ and $l \leq m_1 - 3$, $p'_1, \ldots, p'_{j-1} \in u[1, p_i]$ and $p'_{j+l+1},\ldots, p'_{m_1} \in u[p_{i+1}, |u|]$, then from (3), there exist $q'_j, q'_{j+1}, \ldots, q'_{j+1} \in v(q_i, q_{i+1})$ $(q'_j \leq q'_{j+1} \leq \ldots \leq q'_{j+1})$ such that $Q_a^u p'_r$ if and only if $Q_a^v q'_r$ for $j \leq r \leq j + 1$, $u(p'_r, p'_{r+1}) \sim(m_2)\ v(q'_r, q'_{r+1})$ for $j \leq r \leq j + l - 1$, $u(p_i, p'_j) \sim(m_2)\ v(q_i, q'_j)$ and $u(p'_{j+1}, p_{i+1}) \sim(m_2)\ v(q'_{j+1}, q_{i+1})$.

Case 3. If $p'_j, p'_{j+1}, \ldots, p'_{j+m_1-2} \in u(p_i, p_{i+1})$ and $p'_j < \ldots < p'_{j+m_1-2}$ for some i, $1 \leq i \leq s - 1$ ($j = 1$ and $p'_{m_1} \in u[p_{i+1}, |u|]$) ($j= 2$ and $p'_1 \in u[1, p_i]$ is similar), then from (4), there exist $q'_1, \ldots, q'_{m_1-1} \in v(q_i, q_{i+1})$ $(q'_1 < \ldots < q'_{m_1-1})$ such that (1'), (2') and (3'') hold.

Case 4. If $p'_1, \ldots, p'_{m_1} \in u(p_i, p_{i+1})$ and $p'_1 < \ldots < p'_{m_1}$ for some i, $1 \leq i \leq s - 1$, then from (5), there exist $q'_1, \ldots, q'_{m_1} \in v(q_i, q_{i+1})$ $(q'_1 < \ldots < q'_{m_1})$ such that (1''') and (2''') hold.

From the choice of the p_i, q_i and $q'_1, \ldots, q'_{m_1}$, $q'_1, \ldots, q'_{m_1} \in v$ are such that $q'_1 \leq \ldots \leq q'_{m_1}$, $Q_a^u p'_j$ if and only if $Q_a^v q'_j$, $a \in A$ for $1 \leq j \leq m_1$, $u[1, p'_1) \sim(m_2)\ v[1, q'_1)$, $u(p'_j, p'_{j+1}) \sim(m_2)\ v(q'_j, q'_{j+1})$ for $1 \leq j \leq m_1 - 1$ and $u(p'_{m_1}, |u|] \sim(m_2)\ v(q'_{m_1}, |v|]$. By Lemma 2.1, $u \sim(m_1, m_2)\ v$.

Conversely, assume $u \sim(m_1, m_2)\, v$. (1) and (2) obviously hold. (3) holds. To see this, let $p'_1, \dots, p'_{m_1-2}$ ($p'_1 \leq \dots \leq p'_{m_1-2}$) in $u(p_i, p_{i+1})$ (the proof is similar when starting with $q'_1, \dots, q'_{m_1-2}$ in $v(q_i, q_{i+1})$). Consider the following play of the game $\mathcal{G}(m_1, m_2)(u, v)$. Player I, in the first move, chooses $p_i, p'_1, \dots, p'_{m_1-2}, p_{i+1}$. Hence there exist $q'_1, \dots, q'_{m_1-2}$ ($q'_1 \leq \dots \leq q'_{m_1-2}$) in $v(q_i, q_{i+1})$ such that $u(p_i, p'_1) \sim(m_2)\, v(q_i, q'_1)$, $u(p'_j, q'_{j+1}) \sim(m_2)\, v(q'_j, q'_{j+1})$ for $1 \leq j \leq m_1 - 3$ and $u(p'_{m_1-2}, p_{i+1}) \sim(m_2)\, v(q'_{m_1-2}, q_{i+1})$ (note that from the choice of the p_i and the q_i, $q'_1, \dots, q'_{m_1-2}$ must be in $v(q_i, q_{i+1})$) (4) and (5) similarly follow.

We are now interested in the M-varieties V_{2,m_1} for $m_1 > 1$. For $m_2 \geq 1$, $\mathcal{C}^1_{(m_1,m_2)}$ is a class of equations consisting of

$$u_1 \dots u_{m_2}(xy)^{m_1+(m_1-1)m_2}(xy)(yx)^{m_1+(m_1-1)m_2}v_1 \dots v_{m_2}$$
$$= u_1 \dots u_{m_2}(xy)^{m_1+(m_1-1)m_2}(yx)(yx)^{m_1+(m_1-1)m_2}v_1 \dots v_{m_2}$$

where the u and the v are of the form $x^e y$, $y^e x$, xy^e or yx^e for some e, $1 \leq e \leq \mathcal{N}(m_1, m_2)$. The equation

$$(xy)^{m_2}(xy)^{m_1+(m_1-1)m_2}(xy)(yx)^{m_1+(m_1-1)m_2}(xy)^{m_2}$$
$$= (xy)^{m_2}(xy)^{m_1+(m_1-1)m_2}(yx)(yx)^{m_1+(m_1-1)m_2}(xy)^{m_2}$$

is an example.

$\mathcal{C}^2_{(m_1,m_2)}$ consists of the equations

$$u_1 \dots u_i x^{m_2-i} x^{m_1+(m_1-1)m_2} x^{m_2-j} v_1 \dots v_j$$
$$= ul \dots u_i x^{m_2-i} x^{m_1+(m_1-1)m_2+1} x^{m_2-j} v_1 \dots v_j$$

where the u and the v are as above and $0 \leq i, j \leq m_2$. The equation

$$(xy)^{m_2} x^{m_1+(m_1-1)m_2}(xy)^{m_2} = (xy)^{m_2} x^{m_1+(m_1-1)m_2+1}(xy)^{m_2}$$

is an example.

Theorem 3.5. *Let $m_1 \geq 1$. Every monoid in V_{2,m_1} satisfies $\mathcal{C}^1_{(m_1,m_2)} \cup \mathcal{C}^2_{(m_1,m_2)}$ for all sufficiently large m_2.*

Proof. Similar to Lemma 3.3 using the preceding lemma. The power $m_1 + (m_1—l)m_2$ in $\mathcal{C}^2_{(m_1,m_2)}$ comes from condition (5) in Lemma 3.4.

4. On the dot-depth of the $A^* / \sim(m_1, \dots, m_k)$

Let A contain at least two letters. Let $k \geq 1$. Let $m_1, \dots, m_k$ be positive integers. We are interested in the problem of finding necessary and sufficient conditions for $A^* / \sim(m_1, \dots, m_k)$ to be of dot-depth exactly d. This section gives conditions on $m_1, \dots, m_k$ for d=1 and d=2. Also, we show that for $k \geq 3$, and $m_i \geq 2$ for $2 \leq i \leq k - 1$, $A^*/ \sim(m_1,...,m_k)$ is of dot-depth exactly k. Other results about upper bounds and lower bounds are also discussed.

Theorem 4.1. *Let $k \geq 1$. Let $m_1, \dots, m_k$ be positive integers.* $A^*/ \sim(m_1, ...,m_k)$ *is of dot-depth exactly* 1 *if and only if* k =1.

Proof. We show that for $m_1, m_2 \geq 1$, there is no $m > 0$ such that $A^*/ \sim(m_1, m_2)$ satisfies the equation $u_m = (xy)^m = (yx)^m = v_m$, where x and y are arbitrary distinct letters. We illustrate a winning strategy for player I. (I, i) ((II, i)) denotes a position chosen by player I (II) in the ith move, I = 1,2. Let $N \geq \mathcal{N}(m_1, m_2)$.

$u_N = \dots (xy)(xy)$
(II,1)↑↑(I,2)
$v_N = \dots (yx)(yx)$
↑(I,1)

Player I, in the first move, chooses the last x in v_N. Player II, in the first move, has to choose the last x in u_N (if not, player I in the second move could win by choosing the last x in u_N). Player I, in the second move, chooses the last y in u_N. Player II, in the second move, cannot choose a y in v_N to the right of the previously chosen position in v_N. Hence II loses.

Theorem 4.2. *Let $k \geq 3$. Let m_i, $1 \leq i \leq k$ be positive integers and $m_i \geq 2$ for $2 \leq i \leq k - 1$. Then $A^*/\sim(m_1,...,m_k)$ is of dot-depth exactly k.*

Proof. Let $m>0$. Consider $u_m = (x^{(k-1)}y^{(k-1)})^m$, $v_m = (y^{(k-1)}x^{(k-1)})^m$ (here, $x^{(1)} = x$, $y^{(1)} = y$ and $x^{(r+1)} = (x^{(r)}y^{(r)})^m x^{(r)}(x^{(r)}y^{(r)})^m$, $y^{(r+1)} = (x^{(r)}y^{(r)})^m y^{(r)}(x^{(r)}y^{(r)})^m$). A result of Straubing [191 implies that monoids in V_{k-1} satisfy $u_m = v_m$ for all sufficiently large m. However, for every $N \geq \mathcal{N}(1,2, ...,2,1)$ where $(1,2, \ldots, 2, 1)$ is a k- tuple, $u_N \nsim(1,2, ...,2,1)\, v_N$. A winning strategy for player I in the game $\mathcal{G}(1, 2, ...,2,1)(u_N, v_N)$ is as follows. (I, i) ((II, i)) denotes a position chosen by player I (II) in the ith move, $i = 1, \ldots, k$. Let $N \geq \mathcal{N}(m_1, \ldots, m_k)$. Using $x^N \sim(m_1, \ldots, m_k)\, x^{N+1}$ (Section 2), one sees that

$$u_N \sim(m_1, \ldots, m_k) \ldots (x^{(k-2)}y^{(k-2)})^N \underset{\substack{\uparrow \\ (\mathrm{II},1)}}{x}$$

$$(x^{(k-3)}y^{(k-3)})^N x (x^{(k-3)}y^{(k-3)})^N y (x^{(k-3)}y^{(k-3)})^N (x^{(k-2)}y^{(k-2)})^{N-2}$$

$$(x^{(k-3)}y^{k-3)})^N x (x^{(k-3)}y^{(k-3)})^N \underset{\uparrow}{y} (x^{(k-3)}y^{(k-3)})^N \underset{\uparrow}{y} (x^{(k-2)}y^{(k-2)})^N$$

$$(\mathrm{I},2)$$

Similarly,

$$v_N \sim(m_1, \ldots, m_k) \ldots (x^{(k-2)}y^{(k-2)})^N \underset{\substack{\uparrow \\ (\mathrm{I},1)}}{x}$$

$$(x^{(k-3)}y^{(k-3)})^N x (x^{(k-3)}y^{(k-3)})^N y (x^{(k-3)}y^{(k-3)})^N$$

$$(x^{(k-2)}y^{(k-2)})^{M_1-1} (x^{(k-3)}y^{(k-3)})^N x (x^{(k-3)}y^{(k-3)})^N$$

$$\underset{\uparrow}{y} (x^{(k-3)}y^{(k-3)})^N x (x^{(k-3)}y^{(k-3)})^N \underset{\uparrow}{y} (x^{(k-3)}y^{(k-3)})^N (x^{(k-2)}y^{(k-2)})^{M_2}$$

$$(\mathrm{II},2)$$

where $M_1 + M_2 = N - 2$. Player I, in the first move, chooses the middle x of the last $x^{(k-2)}$ followed immediately by an $x^{(k-2)}$ in v_N. Player II, in the first move, has to choose the middle x of the last $x^{(k-2)}$ followed immediately by an $x^{(k-2)}$ in u_N (if not, player I in the next k-1 moves could win by choosing in the second move the middle x of the last two consecutive $x^{(k-2)}$ in u_N). Player I, in the second move, chooses the middle y of the last two consecutive $y^{(k-2)}$ in u_N. Player II, in the second move, cannot choose the middle y of the last two consecutive $y^{(k-2)}$ in v_N to the right of the previously chosen position. Hence he is forced to choose two $y^{(k-2)}$ by an $x^{(k-2)}$. Player I, in the third move, chooses the middle x of the last two consecutive $x^{(k-3)}$ in v_N between the positions chosen in the preceding move by II. Player II, in the third move, cannot chose the middle x of the last two consecutive $x^{(k-3)}$ in u_N between the previously choosen position by I. Hence he is forced to choose two $x^{(k-3)}$ separated by an $y^{(k-3)}$ and so on. Player I, in the $(k-1)$th move, chooses the last two consecutive x (or y) in v_N (or u_N) between the chosen positions in the preceding move by II. Player II, in the $(k - 1)$th move, is forced to choose two x (or y) in u_N (or v_N) separated by a y (or an x). Player I, in the last move, selects that y (or x). Player II loses since he cannot choose a y (or x) between the two consecutive x chosen in the $(k-1)$th move by I. The result follows.

Note that the infinity of the Straubing hierarchy for an alphabet of at least two letters follows from the preceding theorem.

Theorem 4.3. *Let $k \geq 2$ and d be the dot-depth of $A^*/\sim(m_1, \ldots, m_{2k-2})$. Then $k \leq d \leq 2k-2$.*

Proof. For $k \geq 3$, the upper bound follows from the congruence characterization of V_{2k-2}. Now, by Lemma 2.2,

$$\sim(1, \underbrace{1, \ldots, 1,}_{2k-4} 1) \subseteq (1, \underbrace{2, \ldots, 2,}_{k-2} 1).$$

If

$$A^*/\sim(\underbrace{1, \ldots, 1}_{2k-2})$$

is of dot-depth $<k$, then

$$A^*/\sim(1, \underbrace{2, \ldots, 2}_{k-2}, 1)$$

is also of dot-depth $<k$ since

$$A^*/\sim(1, \underbrace{2, \ldots, 2}_{k-2}, 1) < A^*/\sim(\underbrace{1, \ldots, 1}_{2k-2}).$$

But by Theorem 4.3,

$$A^*/\sim(1, \underbrace{2, \ldots, 2}_{k-2}, 1)$$

is of dot-depth k. For $k = 2$, the result follows from Theorem 4.1.

Theorem 4.4. *Let $k \geq 1$. Let $m_1, \ldots, m_k$ be positive integers. $A^*/\sim(m_1,\ldots,m_k)$ is of dot-depth exactly 2 if and only if*

(1) $k = 2$ or
(2) $k = 3$ and $m_2=1$.

Proof. A result of [1] states that $A^*/\sim(m_1, m_2, m_3)$ is of dot-depth exactly 2 if and only if $m_2= 1$. The theorem follows from that result, Theorems 4.1 and 4.3 and the fact that $u_N = (x^{(2)}y^{(2)})^N$, $v_N = (y^{(2)}x^{(2)})^N$ for $N \geq \mathcal{N}(1, 2, 1)$ in Theorem 4.2 are such that $u_N = (1,2, 1)v_N$ and hence $u_N \nsim (1,1, 1, 1)v_N$ by Lemma 2.2.

Other upper and lower bounds results follow for monoids like $A^*/\sim(1, 1,1,2,1)$. Since $\sim(1,1,1,2, 1) \subseteq \sim(1,3,2, 1)$ by Lemma 2.2, and $A^*/\sim(1,3,2,1)$ is of dot-depth exactly 4 by Theorem 4.2, $A^*/\sim(1, 1, 1,2, 1)$ is of dot-depth $\geq$ 4 and ≤ 5. Similarly, for $A^*/\sim(1,2, 1, 1, 1)$,

References

[1] F. Blanchet-Sadri, On dot-depth two, RAIRO Inform. Theor. Appl. 24 (1990) 521-530.
[2] F. Blanchet-Sadri, Some logical characterizations of the dot-depth hierarchy and applications, Tech. Rept. 88-03, Department of Mathematics and Statistics, McGill University, Montreal, Que. (1988) 1-44.
[3] F. Blanchet-Sadri, Games, equations and the dot-depth hierarchy, Comput. Math. Appl. 18 (1989) 809-822.
[4] L.A. Brzozowski and R. Knast, The dot-depth hierarchy of star-free languages is infinite, J. Comput. System Sci. 16 (1978) 37-55.
[5] L.A. Brzozowski and 1. Simon, Characterizations of locally testable events, Discrete Math. 4 (1973) 243-271.
[6] R.S. Cohen and J.A. Brzozowski, Dot-depth of star-free events, J. Comput. System Sci. 5 (1971) 1-16.
[7] A. Ehrenfeucht, An application of games to the completeness problem for formalized theories, Fund. Math. 49 (1961) 129-141.
[8] S. Eilenberg, Automata, Languages and Machines Vol. B (Academic Press, New York, 1976).
[9] H.B. Enderton, A Mathematical Introduction to Logic (Academic Press, New York, 1972).
[10] R. Knast, A semigroup characterization of dot-depth one languages, RAIRO Inform. Theor. 17 (1983) 321-330.
[1 1] R. Knast, Some theorems on graph congruences, RAIRO Inform. Theor. 17 (1983) 331-342.
[12] R. McNaughton and S. Papert, Counter-Free Automata (MIT Press, Cambridge, MA, 1971).
[13] D. Perrin and J.E. Pin, First order logic and star-free sets, J. Comput. System Sci. 32 (1986) 393-406.
[14] T.E. Pin, Varietes de Languages Formels (Masson, Paris, 1984).

[151 J.E. Pin, Hierarchies de concatenation, RAIRO Inform. Theor. 18 (1984) 23-46. [16] .1 G. Rosenstein, Linear Orderings (Academic Press, New York, 1982).
[171 M.P. Schritzenberger, On finite monoids having only trivial subgroups, Inform. and Control 8 (1965) 190-194.
[18] 1. Simon, Piecewise testable events, in: Proceedings of the 2nd GI Conference, Lecture Notes in Computer Science 33 (Springer, Berlin, 1975) 214-222.
[19] H. Straubing, A generalization of the Schritzenberger product of finite monoids, Theoret. Comput. Sci. 13 (1981) 137-150.
[20] H. Straubing, Finite semigroup varieties of the form V `D, J. Pure Appl. Algebra 36 (1985) 53-94.
[21] H. Straubing, Semigroups and languages of dot-depth two, in: Proceedings of the 13th !CALF, Lecture Notes in Computer Science 226 (Springer, New York, 1986) 416-423.
122] W. Thomas, Classifying regular events in symbolic logic, J. Comput. System Sci. 25(1982) 360-176.
[23] W. Thomas, An application of the Ehrenfeucht-Fraisse game in formal language theory, Bull. SOC Math. France 16 (1984) 11-21.
[24] B. Tilson, Categories as algebra, J. Pure Appl. Algebra 48 (1987) 83-198.