Electronic Colloquium on Computational Complexity, Report No. 40 (2013)

The Algebraic Theory of Parikh Automata

Michaël Cadilhac[1], Andreas Krebs[2], and Pierre McKenzie[1,⋆]

[1] DIRO, Université de Montréal
C.P. 6128 succ. Centre-Ville, Montréal (Québec), H3C 3J7 Canada
{cadilhac, mckenzie}@iro.umontreal.ca
[2] Wilhelm-Schickard-Institute, Universität Tübingen
Tübingen, Germany
krebs@informatik.uni-tuebingen.de

Abstract. The Parikh automaton model equips a finite automaton with integer registers and imposes a semilinear constraint on the set of their final settings. Here the theory of typed monoids is used to characterize the language classes that arise algebraically. Complexity bounds are derived, such as containment of the unambiguous Parikh automata languages in NC^1. Noting that DetAPA languages are positive supports of rational $\mathbb{Z}$-series, DetAPA are further shown stronger than Parikh automata on unary langages. This suggests unary DetAPA languages as candidates for separating the two better known variants of uniform NC^1.

Introduction

The Parikh automaton model was introduced in [1]. It amounts to a nondeterministic finite automaton equipped with registers tallying up the number of occurrences of each transition along an accepting run. Such a run is then deemed successful iff the tuple of final register settings falls within a fixed semilinear set. An *affine* variant of the model in which transitions further induce an affine transformation on the registers was considered in [2]. An *unambiguous* variant of the model was considered in [3]. Tree Parikh automata and other variants were considered in [4].

Recall the tight connection between AC^0, ACC^0 and NC^1 and aperiodic monoids, solvable monoids and nonsolvable monoids respectively [5,6]. This connection was refined and studied in depth (see [7] for a lovely account), but the class $\mathrm{TC}^0 \subseteq \mathrm{NC}^1$ was left out of the picture because the MAJ gate in circuits could not be translated into the operation of a finite algebraic structure. Typed monoids were introduced in [8] as a means of capturing TC^0 meaningfully in the algebraic framework.

In both the classical and the typed monoid framework, a compelling notion of a natural class of monoids is that of a variety. In both frameworks, different monoid varieties capture different classes of languages as inverse homomorphic images of an accepting subset of the monoid [9,10]. The internal structure of

⋆ Supported by the Natural Sciences and Engineering Research Council of Canada

ISSN 1433-8092

NC^1 hinges on whether different monoid varieties still capture different classes of languages when the classical notion of a homomorphism is appropriately generalized to capture as above complexity classes such as ACC^0, TC^0 and NC^1.

Our contribution is an algebraic characterization of the language classes defined by the deterministic and unambiguous variants of the Parikh automaton (called CA, for "constrained automaton") and the affine Parikh automaton. We show:

- the class $\mathcal{L}_{\mathrm{DetCA}}$ of languages accepted by deterministic CA is the set of languages recognized by typed monoids from $\mathbf{Z}^+ \wr \mathbf{M}$, i.e., by wreath products of the monoid of integers with some finite monoid; the least typed monoid variety generated by $\mathbf{Z}^+ \wr \mathbf{M}$ also captures $\mathcal{L}_{\mathrm{DetCA}}$
- the class $\mathcal{L}_{\mathrm{UnCA}}$ of languages accepted by unambiguous CA is the set of languages recognized by typed monoids from $\mathbf{Z}^+ \Box \mathbf{M}$, i.e., by block products of the monoid of integers with some finite monoid; the least typed monoid variety generated by $\mathbf{Z}^+ \Box \mathbf{M}$ also captures $\mathcal{L}_{\mathrm{UnCA}}$
- the classes $\mathcal{L}_{\mathrm{DetAPA}}$ and $\mathcal{L}_{\mathrm{UnAPA}}$, of languages accepted by deterministic and by unambiguous affine Parikh automata respectively (where an affine Parikh generalizes the constrained automaton by allowing each transition to perform an affine transformation on the automaton registers), are the Boolean closure of the positive supports of rational series over the integers.

The first two characterizations above add legitimacy to the theory of typed monoids, and they suggest further relevance of that theory to our understanding of NC^1. It follows from the characterization of $\mathcal{L}_{\mathrm{UnCA}}$ that $\mathcal{L}_{\mathrm{UnCA}} \subseteq NC^1$, a fact which is not immediately obvious from the operation of an unambiguous constrained automaton.

The Boolean closure of the class of positive supports of rational series over the integers, hence $\mathcal{L}_{\mathrm{DetAPA}} = \mathcal{L}_{\mathrm{UnAPA}}$, can be viewed as a very tightly uniform version of the (DLOGTIME-uniform) class PNC^1, introduced in [11] as the log depth analog of the poly time and log space classes PP and PL [12]. Fulfilling $NC^1 \subseteq PNC^1 \subseteq L$, PNC^1 is robust, pointedly characterized using iterated products of constant dimension integer matrices, but also characterized using paths in bounded width graphs, proof trees in log depth circuits, accepting paths in nondeterministic finite automata or evaluation of a log depth $\{+, \times\}$-formula [11]. An elaborate structural complexity evolved around PNC^1 with the work of [13]. We note that using formal power series as a tool to investigate counting classes below L was already suggested in [14], but with emphasis there on the complexity of performing operations such as inversion and root extraction on such series.

1 Preliminaries

Integers, vectors, monoids We write $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Z}^+$, $\mathbb{Z}_0^-$ for the sets of nonnegative integers, integers, positive integers, and nonpositive integers respectively. Vectors in $\mathbb{N}^d$ are noted in bold, e.g., $\mathbf{v}$ whose elements are $v_1, v_2, \ldots, v_d$. We write

$\mathbf{e}_i \in \{0,1\}^d$ for the vector having a 1 only in position i and $\mathbf{0}$ for the all-zero vector. We view $\mathbb{N}^d$ as the additive monoid $(\mathbb{N}^d, +)$, with $+$ the component-wise addition and $\mathbf{0}$ the identity element. We let $\mathcal{M}_{\mathbb{Z}}(k)$, for $k \geq 1$, be the monoid of square matrices of dimension $k \times k$ with values in $\mathbb{Z}$ and with the operation being the inverted matrix multiplication. We write Ψ_i for the projection on the i-th component.

For a monoid M, we write e_M for the identity element of M. For $S \subseteq M$, we write S^* for the monoid generated by S, i.e., the smallest submonoid of M containing S. A (monoid) *morphism* from $(M, \cdot)$ to $(N, \circ)$ is a function $h\colon M \to N$ such that $h(m_1 \cdot m_2) = h(m_1) \circ h(m_2)$, and $h(e_M) = e_N$. Moreover, if $M = S^*$ for some finite set of symbols S (and this will always be the case), then h need only be defined on the elements of S. In this case, h is said to be *erasing* if there is an $s \in S$ such that $h(s) = e_N$. If in addition $N = \mathrm{T}^*$ for some finite set of symbols T, h is said to be *length-preserving* if for all $s \in S$, $h(s) \in \mathrm{T}$. Also, note that a morphism $h\colon \{a,b\}^* \to \mathcal{M}_{\mathbb{Z}}(k)$ is such that $h(ab) = h(b).h(a)$ with . the usual matrix multiplication.

Semilinear sets, Parikh image A subset C of $\mathbb{N}^d$ is *linear* if there exist $\mathbf{c} \in \mathbb{N}^d$ and a finite $P \subseteq \mathbb{N}^d$ such that $C = \mathbf{c} + P^*$. The subset C is said to be *semilinear* if it is equal to a finite union of linear sets: $\{4n + 56 \mid n > 0\}$ is semilinear while $\{2^n \mid n > 0\}$ is not. We will often use the fact that the semilinear sets are those sets of natural numbers definable in first-order logic with addition [15]. Let $\Sigma = \{a_1, a_2, \ldots, a_n\}$ be an (ordered) alphabet with ε the empty word. The *Parikh image* is the morphism $\mathsf{Pkh}\colon \Sigma^* \to \mathbb{N}^n$ defined by $\mathsf{Pkh}(a_i) = \mathbf{e}_i$, for $1 \leq i \leq n$ — in particular, we have that $\mathsf{Pkh}(\varepsilon) = \mathbf{0}$. For $w \in \Sigma^*$, with $\mathsf{Pkh}(w) = \mathbf{x}$ and $a \in \Sigma$, we write $|w|_a$ for x_a. The Parikh image of a language L is defined as $\mathsf{Pkh}(L) = \{\mathsf{Pkh}(w) \mid w \in L\}$. The name of this morphism stems from Parikh's theorem [16], stating that for L context-free, $\mathsf{Pkh}(L)$ is semilinear.

Affine Functions A function $f\colon \mathbb{N}^d \to \mathbb{N}^d$ is a (total and positive) *affine function* of dimension d if there exist a matrix $M \in \mathbb{N}^{d \times d}$ and $\mathbf{v} \in \mathbb{N}^d$ such that for any $\mathbf{x} \in \mathbb{N}^d$, $f(\mathbf{x}) = M.\mathbf{x} + \mathbf{v}$. We abusively write $f = (M, \mathbf{v})$. We let $\mathcal{F}_d$ be the monoid of such functions under the operation $\diamond$ defined by $(f \diamond g)(\mathbf{x}) = g(f(\mathbf{x}))$, where the identity element is the identity function, i.e., $(Id, \mathbf{0})$ with Id the identity matrix of dimension d. Let U be a monoid morphism from Σ^* to $\mathcal{F}_d$. For $w \in \Sigma^*$, we write U_w for $U(w)$, so that the application of $U(w)$ to a vector $\mathbf{v}$ is written $U_w(\mathbf{v})$, and U_ε is the identity function.

Automata An automaton is a quintuple $A = (Q, \Sigma, \delta, q_0, F)$ where Q is a finite set of states, Σ is an alphabet, $\delta \subseteq Q \times \Sigma \times Q$ is a set of transitions, $q_0 \in Q$ is the initial state, and $F \subseteq Q$ is a set of final states. For a transition $t = (q, a, q') \in \delta$, we write $t = q \bullet\!\!- a \!\to q'$ and define $\mathsf{From}(t) = q$ and $\mathsf{To}(t) = q'$. We define $\mu_A\colon \delta^* \to \Sigma^*$ as the length-preserving morphism given by $\mu_A(t) = a$, with, in particular, $\mu_A(\varepsilon) = \varepsilon$, and write μ when A is clear from the context. The set of accepting paths of A, i.e., the set of words over δ describing paths starting from q_0 and ending in F, is written $\mathsf{Run}(A)$. The language of the automaton is

$L(A) = \mu_A(\mathsf{Run}(A))$. An automaton is *unambiguous* if for all $w \in L(A)$ there is a unique $\pi \in \mathsf{Run}(A)$ with $\mu(\pi) = w$.

A *constrained automaton (CA)* [2] is a pair (A, C) where A is an automaton with d transitions and $C \subseteq \mathbb{N}^d$ is semilinear. Its language is $L(A, C) = \mu_A(\{\pi \in \mathsf{Run}(A) \mid \mathsf{Pkh}(\pi) \in C\})$. The CA is said to be *deterministic (DetCA)* if A is deterministic, and *unambiguous (UnCA)* if A is unambiguous. We write $\mathcal{L}_{\text{CA}}$, $\mathcal{L}_{\text{DetCA}}$, and $\mathcal{L}_{\text{UnCA}}$ for the classes of languages recognized by CA, DetCA, and UnCA, respectively.

An *affine Parikh automaton (APA)* of dimension d is a triple (A, U, C) where A is an automaton with transition set δ, $U : \delta^* \to \mathcal{F}_d$ is a morphism, and $C \subseteq \mathbb{N}^d$ is semilinear. Its language is $L(A, U, C) = \mu_A(\{\pi \in \mathsf{Run}(A) \mid U_\pi(\mathbf{0}) \in C\})$. The APA is said to be *deterministic (DetAPA)* if A is deterministic, and *unambiguous (UnAPA)* if A is unambiguous. We write $\mathcal{L}_{\text{APA}}$, $\mathcal{L}_{\text{DetAPA}}$, and $\mathcal{L}_{\text{UnAPA}}$ for the classes of languages recognized by APA, DetAPA, and UnAPA, respectively.

Transition monoid Let $A = (Q, \Sigma, \delta, q_0, F)$ be a complete deterministic automaton. For $a \in \Sigma$, define $f_a : Q \to Q$ by $f_a(q) = q'$ iff $q \bullet\!\!-a\!\!\to q' \in \delta$. The *transition monoid* M of A is the closure under composition of the set $\{f_a \mid a \in \Sigma\}$. The monoid M acts on Q naturally by $q^m = m(q)$, $m \in M$, $q \in Q$. Write $\eta : \Sigma^* \to M$ for the canonical surjective morphism associated, that is, the morphism defined by $\eta(a) = f_a$, $a \in \Sigma$. Then $q^{\eta(w)}$ is the state reached by reading $w \in \Sigma^*$ from the state $q \in Q$.

2 Normal forms of CA and APA

We present several technical lemmata on CA and APA, that will help us in devising concise proofs for the algebraic characterizations that follow. Their main purpose is to simplify the constraint set, so that only sign checks on linear combinations of variables are performed.

Recall (e.g., [17]) that for any semilinear set $C \subseteq \mathbb{Z}^d$, there is a Boolean combination of expressions of the form: $\sum_{1 \le i \le d} \alpha_i x_i > c$ and $\sum_{1 \le i \le d} \alpha_i x_i \equiv_p c$, with $\alpha_i, c \in \mathbb{Z}$ and $p > 1$, which is true iff $(x_1, x_2, \ldots, x_d) \in C$. Note that the α_i may be zero. We define two notions which refine this point of view:

Definition 1. *We say that a semilinear set C is* modulo-free *if it can be expressed as a Boolean combination of expressions of the form $\sum_i \alpha_i x_i > c$, for $\alpha_i \in \mathbb{Z}$. We say that C is* basic *if it can further be expressed as a* positive *Boolean combination of expressions of the form $\sum_i \alpha_i x_i > 0$.*

The first normal form concerns DetCA and UnCA:

Lemma 1. *Every DetCA (resp. UnCA) has the same language $L \subseteq \Sigma^+$ as another DetCA (resp. UnCA) (A, C) with $L(A) = \Sigma^*$ and C a basic set.*

We also note the following simple fact:

Lemma 2. *For $(A, C_1 \cap C_2)$ a DetCA or an UnCA it holds that:*

$$L(A, C_1 \cap C_2) = L(A, C_1) \cap L(A, C_2) \ .$$

The same holds for $\cup$.

We show more in the context of APA to allow the forthcoming proofs of characterization to translate smoothly from CA to APA. In the following, we consider that a matrix $M \in \mathcal{M}_{\mathbb{Z}}(k)$ is in a set $C \subseteq \mathbb{Z}^{k^2}$ if the vector consisting of the columns of M is in C.

Lemma 3. *Let $L \subseteq \Sigma^+$ be in $\mathcal{L}_{\mathrm{DetAPA}}$. There is a morphism $h\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(k)$, for some k, and a set $\mathcal{Z} \subseteq \mathbb{Z}^{k^2}$ expressible as a Boolean combination of expressions $x_i > 0$, such that $L = h^{-1}(\mathcal{Z})$.*

Similarly, let $L \subseteq \Sigma^+$ be in $\mathcal{L}_{\mathrm{APA}}$ (resp. in $\mathcal{L}_{\mathrm{UnAPA}}$). There is an automaton (resp. unambiguous automaton) A with transition set δ, a morphism $h\colon \delta^* \to \mathcal{M}_{\mathbb{Z}}(k)$, for some k, and a set $\mathcal{Z} \subseteq \mathbb{Z}^{k^2}$ expressible as a Boolean combination of expressions $x_i > 0$, such that $L = \mu_A(h^{-1}(\mathcal{Z}) \cap \mathsf{Run}(A))$.

3 Capturing Parikh automata classes algebraically

3.1 Typed monoids

In this section we characterize DetCA, UnCA, DetAPA, and UnAPA using the theory of (finitely) typed monoids [8].

Definition 2 (Typed monoid [8]). *A* typed monoid *is a pair $(S, \mathfrak{S})$ where S is a finitely generated monoid and $\mathfrak{S}$ is a finite Boolean algebra of subsets of S whose elements are called* types. *We write $(S, \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n\})$ for the typed monoid $(S, \mathfrak{S})$ where $\mathfrak{S}$ is generated by the $\mathcal{S}_i$'s. If $n = 1$, we simply write $(S, \mathcal{S}_1)$. For two typed monoids $(M, \mathfrak{M})$, $(N, \mathfrak{N})$, their direct product $(S, \mathfrak{S}) = (M, \mathfrak{M}) \times (N, \mathfrak{N})$ is defined by $S = M \times N$, and $\mathfrak{S}$ is the Boolean algebra generated by $\{\mathcal{M} \times \mathcal{N} \mid \mathcal{M} \in \mathfrak{M} \wedge \mathcal{N} \in \mathfrak{N}\}$. A typed monoid $(S, \mathfrak{S})$* recognizes *a language L if there are a morphism $h\colon \Sigma^* \to S$ and a type $\mathcal{S} \in \mathfrak{S}$ such that $L = h^{-1}(\mathcal{S})$. We write $\mathcal{L}((S, \mathfrak{S}))$ for the class of languages, over any alphabet, recognized by $(S, \mathfrak{S})$ and extend this notation naturally to classes of typed monoids.*

We view a finite monoid M as the typed monoid $(M, 2^M)$, and write $\mathbf{M}$ for the class of typed finite monoids; note that the usual notion of language recognition then coincides with the one given here.

We use the vocabulary of [10] related to typed monoids and recall the main notions in the appendix for completeness.

The appropriateness of typed monoids in the study of the algebraic properties of nonregular languages is witnessed by the following Eilenberg-like theorem of Behle, Krebs, and Reifferscheid:

Theorem 1 ([10]). *Varieties of typed monoids and varieties of languages are in a one-to-one correspondence, i.e., (1) Let $\mathcal{V}$ be a variety of languages and $\mathbf{V}$ the smallest variety of typed monoids that recognizes all languages in $\mathcal{V}$, then $\mathcal{L}(\mathbf{V}) = \mathcal{V}$; (2) Let $\mathbf{V}$ be a variety of typed monoids and $\mathbf{W}$ be the smallest variety that recognizes all languages of $\mathcal{L}(\mathbf{V})$, then $\mathbf{V} = \mathbf{W}$.*

Similar to the untyped algebraic theory of languages, if a typed monoid recognizes a language, it also recognizes its complement. This implies that $\mathcal{L}_{\mathrm{CA}}$, which is not closed under complement, does not accept a typed monoid characterization. We will thus focus on characterizing the deterministic and unambiguous classes. Note that we will frequently focus on languages which do not contain the empty word. This is a technical simplification which introduces no loss of generality, as all our typed monoid classes recognize $\{\varepsilon\}$ and are closed under union.

3.2 Capturing DetCA and UnCA

Let $\mathbf{Z}^+$ be the set of typed monoids $\{(\mathbb{Z}, \mathbb{Z}^+)^k \mid k \geq 1\}$.

Theorem 2. $\mathcal{L}(\mathbf{Z}^+ \wr \mathbf{M}) = \mathcal{L}_{\mathrm{DetCA}}$*, where* $\wr$ *denotes the wreath product.*

Proof. ($\mathcal{L}_{\mathrm{DetCA}} \subseteq \mathcal{L}(\mathbf{Z}^+ \wr \mathbf{M})$) We first show that $\mathcal{L}(\mathbf{Z}^+ \wr \mathbf{M})$ is closed under union and intersection. Let $L_1, L_2 \in \mathcal{L}(\mathbf{Z}^+ \wr \mathbf{M})$ be two languages over Σ, that is, for $i = 1, 2$, there exist a finite monoid M_i, an integer k_i, a morphism $h_i \colon \Sigma^* \to \mathbb{Z}^{k_i} \wr M_i$, and a type $\mathcal{T}_i$ of $(\mathbb{Z}, \mathbb{Z}^+)^{k_i}$ such that $L_i = h_i^{-1}(\mathcal{T}_i)$.

Consider the typed monoid $(\mathbb{Z}, \mathbb{Z}^+)^{k_1+k_2} \wr (M_1 \times M_2) \in \mathbf{Z}^+ \wr \mathbf{M}$. This monoid recognizes both the intersection and union of L_1 and L_2 as follows. Define $h \colon \Sigma^* \to \mathbb{Z}^{k_1+k_2} \wr (M_1 \times M_2)$ by $h(a) = (f_a, (\Psi_2(h_1(a)), \Psi_2(h_2(a))))$ where $a \in \Sigma$ and $f_a((m_1, m_2)) = ([\Psi_1(h_1(a))](m_1), [\Psi_1(h_2(a))](m_2)) \in \mathbb{Z}^{k_1+k_2}$. This function is type-respecting. Now let $\oslash \in \{\cup, \cap\}$. We define $\mathcal{T}_\oslash = (\mathcal{T}_1 \times \mathbb{Z}^{k_2}) \oslash (\mathbb{Z}^{k_1} \times \mathcal{T}_2)$, and thus $h^{-1}(\mathcal{T}_\oslash) = L_1 \oslash L_2$.

Now let (A, C) be a DetCA with $A = (Q, \Sigma, \delta, q_0, F)$, and suppose (by Lemma 1) that $F = Q$ and that the constraint set is expressed by a positive Boolean combination of clauses of the form $\sum_{t \in \delta} \alpha_t x_t > 0$. Closure of $\mathcal{L}(\mathbf{Z}^+ \wr \mathbf{M})$ under $\cup$ and $\cap$ together with Lemma 2 imply that it is enough to argue the case in which C is defined by a single such clause.

Let M be the transition monoid of A, $\eta \colon \Sigma^* \to M$ the canonical morphism associated, and for $m \in M$, write q^m for the action of m on q (i.e., $q^{\eta(w)}$ is the state reached reading w from the state q in A). We now define $h \colon \Sigma^* \to \mathbb{Z} \wr M$ as follows. Let $\tau \colon M \times \Sigma \to \delta$ be defined by $\tau(m, a) = q_0^m \bullet\!\!-a\!\!\to q_0^{m\eta(a)}$. Then:

$$h(a) = (f_a, \eta(a)), \text{ where } f_a(m) = \alpha_{\tau(m,a)} \enspace .$$

Now for $w = w_1 w_2 \cdots w_n \in \Sigma^*$ and $\pi = \pi_1 \pi_2 \cdots \pi_n$ the unique accepting path in A from q_0 labeled w, we have:

$$h(w) = (f_{w_1} + \eta(w_1) \cdot f_{w_2} + \cdots + \eta(w_1 w_2 \cdots w_{n-1}) \cdot f_{w_n}, \eta(w))$$

$$[\Psi_1(h(w))](\eta(\varepsilon)) = \alpha_{\tau(\eta(\varepsilon), w_1)} + \sum_{i=2}^{n} \alpha_{\tau(\eta(w_1 \cdots w_{i-1}), w_i)} \enspace ,$$

note that $q_0^{\eta(w_1\cdots w_{i-1})}$ is $\mathsf{From}(\pi_{i-1})$ and thus $\tau(\eta(w_1\cdots w_{i-1}), w_i) = \pi_i$, hence:

$$[\Psi_1(h(w))](\eta(\varepsilon)) = \sum_{i=1}^{n} \alpha_{\pi_i} = \sum_{t\in\delta} |\pi|_t \times \alpha_t \enspace .$$

Thus, $\mathsf{Pkh}(\pi) \in C$ iff $[\Psi_1(h(w))](\eta(\varepsilon)) > 0$. Hence with the type $\mathcal{T} = \{(f,m) \in (\mathbb{Z},\mathbb{Z}^+) \wr M \mid f(\eta(\varepsilon)) > 0\}$, which is indeed a type of $(\mathbb{Z},\mathbb{Z}^+) \wr M$, we have that $h^{-1}(\mathcal{T}) = L(A,C)$.

($\mathcal{L}(\mathbf{Z}^+ \wr \mathbf{M}) \subseteq \mathcal{L}_{\mathrm{DetCA}}$) Let $L \subseteq \Sigma^*$ be recognized by $(\mathbb{Z},\mathbb{Z}^+)^k \wr M$ using a type $\mathcal{T}$ and a morphism $h\colon \Sigma^* \to (\mathbb{Z}^k)^M \times M$, and write for convenience $h_i(w) = \Psi_i(h(w))$, $i = 1,2$. Let A be the automaton $(M, \Sigma, \delta, e_M, M)$, where:

$$\delta = \{m \bullet\!\!-a\!\!\rightarrow m' \mid m \in M \wedge a \in \Sigma \wedge m' = m.h_2(a)\} \enspace .$$

Now as $\mathcal{L}_{\mathrm{DetCA}}$ is closed under union and intersection, we may suppose that the type $\mathcal{T}$ is of the following form:

$$\mathcal{T} = \prod_{i=1}^{k} \{(f,m) \mid f(e_M) \in \mathcal{T}_i\} \enspace ,$$

where each $\mathcal{T}_i \in \{\emptyset, \mathbb{Z}_0^-, \mathbb{Z}^+, \mathbb{Z}\}$. Define $T = \mathcal{T}_1 \times \mathcal{T}_2 \times \cdots \times \mathcal{T}_k$, and the semilinear set C consisting of elements:

$$(x_{t_1}, x_{t_2}, \ldots, x_{t_{|\delta|}}) \text{ s.t. } \sum_{t\in\delta} x_t \times [h_1(\mu(t))](\mathsf{From}(t)) \in T \enspace .$$

We claim that the language of the DetCA (A,C) is L. Let $w = w_1 w_2 \cdots w_n \in \Sigma^*$. There is an (accepting) path in A labeled w going through the states $e_M = h_2(\varepsilon)$, $h_2(w_1)$, $h_2(w_1w_2)$, $\ldots$, $h_2(w_1w_2\cdots w_n)$. Thus the sum computed by the semilinear set is $h_1(w_1) + h_2(w_1)\cdot h_1(w_2) + \cdots + h_2(w_1w_2\cdots w_n)\cdot h_1(w_n)$, taken at the point e_M. This is precisely $[h_1(w)](e_M)$, and thus checking whether it belongs to T is equivalent to checking whether $h(w) \in \mathcal{T}$. Hence $L = L(A,C)$. □

Now $\mathcal{L}_{\mathrm{DetCA}}$ is a variety of languages and we may naturally ask whether the smallest variety containing $\mathbf{Z}^+ \wr \mathbf{M}$, which recognizes only the languages of $\mathcal{L}_{\mathrm{DetCA}}$ by Theorem 1, is closed under iterated wreath product. We note this is not the case. Let $U_1 = (\{0,1\}, \times)$, then:

Theorem 3. *There is a language $L \notin \mathcal{L}_{\mathrm{CA}}$ recognized by[3] $U_1 \wr (\mathbb{Z},\mathbb{Z}^+)$ and by $(\mathbb{Z},\mathbb{Z}^+) \wr (\mathbb{Z},\mathbb{Z}^+)$.*

Theorem 4. *$\mathcal{L}(\mathbf{Z}^+ \Box \mathbf{M}) = \mathcal{L}_{\mathrm{UnCA}}$, where $\Box$ denotes the block product.*

[3] The wreath product of two infinite monoids is ill-defined; however Theorem 3 stays true with an adequate definition mimicking that of [8].

Proof. ($\mathcal{L}_{\mathrm{UnCA}} \subseteq \mathcal{L}(\mathbf{Z}^+\Box\mathbf{M})$) We first note that $\mathcal{L}(\mathbf{Z}^+\Box\mathbf{M})$ is closed under union and intersection; this is the same proof as in Theorem 2 except that f_a is now defined as:

$$f_a((m_1, m_2), (m_1', m_2')) = ([\Psi_1(h_1(a))](m_1, m_1'), [\Psi_1(h_2(a))](m_2, m_2')) \enspace .$$

Next consider an UnCA (A, C) with $A = (Q, \Sigma, \delta, q_0, F)$, and suppose (using Lemma 1) that $L(A) = \Sigma^*$ and that the constraint set is expressed by a positive Boolean combination of clauses of the form $\sum_{t\in\delta} \alpha_t x_t > 0$. Closure of $\mathcal{L}(\mathbf{Z}^+\Box\mathbf{M})$ under $\cup$ and $\cap$ together with Lemma 2 imply that it is enough to argue the case in which C is defined by a single such clause.

Let M be the transition monoid of the deterministic version of A, obtained using the powerset construction. Let A' be defined as A with all transitions inverted (i.e., $p \bullet\!\!-a\!\!\to q$ is in A iff $q \bullet\!\!-a\!\!\to p$ is in A'). Let M' be the transition monoid of the deterministic version of A', using again the powerset construction, and let M^{c} be the monoid defined on the same elements as M' but with the operation reversed (i.e., $m_1 \circ_{M'} m_2$ in M' is $m_2 \circ_{M^{\mathrm{c}}} m_1$ in M^{c}; this is still a monoid as $\circ_{M^{\mathrm{c}}}$ is still associative). We will show that $L(A, C)$ is recognized by $(S, \mathfrak{S}) = (\mathbb{Z}, \mathbb{Z}^+)\Box(M \times M^{\mathrm{c}})$.

Write η and η^{c} for the canonical morphisms associated with M and M^{c}; for $m \in M$ and $R \subseteq Q$, write R^m for the action of m on R, and likewise for M^{c}. We first note that for $w \in \Sigma^*$, $\{q_0\}^{\eta(w)}$ is the set of states of A that can be reached in A reading w from q_0, and, likewise, that $F^{\eta^{\mathrm{c}}(w)}$ is the set of states in A from which reading w leads to a final state.

Now for $m_1 \in M$, $a \in \Sigma$, and $m_2 \in M^{\mathrm{c}}$, let $\tau(m_1, a, m_2)$ be the unique transition in A from a state in $\{q_0\}^{m_1}$ to a state in F^{m_2} labeled a. We show that τ is well-defined. Let w_1, w_2 such that $\eta(w_1) = m_1$ and $\eta^{\mathrm{c}}(w_2) = m_2$; this means that there are w_1-labeled paths in A from q_0 to *any* state in $\{q_0\}^{m_1}$, and, likewise, w_2-labeled paths in A from *any* state in F^{m_2} to a final state. *(Existence):* as $w_1 a w_2$ is in $\Sigma^* = L(A)$, there is a transition in A from a state in $\{q_0\}^{m_1}$ to a state in F^{m_2} labeled a. *(Uniqueness):* if two transitions $p \bullet\!\!-a\!\!\to p'$ and $q \bullet\!\!-a\!\!\to q'$ are such that $p, q \in \{q_0\}^{m_1}$ and $p', q' \in F^{m_2}$, this means that there are multiple accepting paths in A labeled $w_1 a w_2$, contradicting the unambiguity of A.

We now define $h\colon \Sigma^* \to S$ by:

$$\begin{aligned} h(a) &= (f_a, (\eta(a), \eta^{\mathrm{c}}(a))), \text{ where} \\ f_a((m_1, m_2), (m_1', m_2')) &= \alpha_{\tau(m_1, a, m_2')} \enspace . \end{aligned}$$

Now let $w = w_1 w_2 \cdots w_i \in \Sigma^*$ and π be the unique path in A from q_0 to a final state labeled w. Then:

$$\begin{aligned} \pi &= \pi_1 \pi_2 \cdots \pi_n \quad \text{where} \\ \pi_i &= \tau(\eta(w_1 w_2 \cdots w_{i-1}), w_i, \eta^{\mathrm{c}}(w_{i+1} w_{i+2} \cdots w_n)) \enspace , \end{aligned}$$

and thus:

$$[\Psi_1(h(w))]((\eta(\varepsilon), \eta^{\mathrm{c}}(\varepsilon))) = \sum_{t\in\delta} |\pi|_t \times \alpha_t \enspace .$$

Thus $\mathsf{Pkh}(\pi) \in C$ iff $[\Psi_1(h(w))]((\eta(\varepsilon), \eta^{\mathrm{c}}(\varepsilon))) > 0$. Hence with the type $\mathcal{S} = \{(f, m) \in S \mid f((\eta(\varepsilon), \eta^{\mathrm{c}}(\varepsilon))) \in \mathbb{Z}^+\}$, which is indeed a type in $\mathfrak{S}$ as $\mathbb{Z}^+$ is a type of $(\mathbb{Z}, \mathbb{Z}^+)$, we have that $h^{-1}(\mathcal{S}) = L(A, C)$.

($\mathcal{L}(\mathbf{Z}^+\Box\mathbf{M}) \subseteq \mathcal{L}_{\mathrm{UnCA}}$) Let $L \subseteq \Sigma^*$ be recognized by $(\mathbb{Z}, \mathbb{Z}^+)^k \Box M$ using a type $\mathcal{T}$ and a morphism $h = (h_1, h_2)$ with $h_1 \colon \Sigma^* \to \mathbb{Z}^{M \times M}$ and $h_2 \colon \Sigma^* \to M$. Let $A(s_1, s_2)$ be the automaton $(M \times M, \Sigma, \delta, (s_1, s_2), M \times \{e_M\})$ where:

$$\begin{aligned}\delta = &\{(m_1, m_2) \bullet\!\!-a\!\!\to (m_1', m_2') \mid \\ &\qquad m_1' = m_1.h_2(a) \wedge h_2(a).m_2' = m_2 \in M \wedge a \in \Sigma\} \ .\end{aligned}$$

Note that $w \in L(A(s_1, s_2))$ implies $h_2(w) = s_2$. We argue that $A(s_1, s_2)$ is unambiguous for any $s_1, s_2 \in M$. It is clear that if $w = \varepsilon$, every $A(s_1, s_2)$ has at most one accepting path labeled w. Now let $w = a \cdot v$ for $v \in \Sigma^*$. Suppose $w \in L(A(s_1, s_2))$. This implies that $h_2(w) = s_2$. The states that can be reached from $(s_1, h_2(w))$ reading a are all of the form $(s_1.h_2(a), m)$, $m \in M$. Now v should be accepted by the automaton A where the initial state is set to one of these states; thus there is only one state fitting, $(s_1.h_2(a), h_2(v))$. By induction hypothesis, there is only one path in $A(s_1.h_2(a), h_2(v))$ recognizing v, thus there is only one path in $A(s_1, h_2(w))$ recognizing w. This shows that for any s_1, s_2, $A(s_1, s_2)$ is unambiguous.

Now, with $e = (e_M, e_M)$, and as $\mathcal{L}_{\mathrm{UnCA}}$ is closed under union and intersection, we may suppose that the type $\mathcal{T}$ is of the following form:

$$\mathcal{T} = \prod_{i=1}^{k} \{(f, m) \mid f(e, e) \in \mathcal{T}_i\} \ ,$$

where each $\mathcal{T}_i \in \{\emptyset, \mathbb{Z}_0^-, \mathbb{Z}^+, \mathbb{Z}\}$. Define $T = \mathcal{T}_1 \times \mathcal{T}_2 \times \cdots \times \mathcal{T}_k$, and the semilinear set C consisting of elements:

$$(x_{t_1}, x_{t_2}, \ldots, x_{t_{|\delta|}}) \text{ s.t. } \sum_{t \in \delta} x_t \times [h_1(\mu(t))](\Psi_1(\mathsf{From}(t)), \Psi_2(\mathsf{To}(t))) \in T \ .$$

We show that $\bigcup_{m \in M} L(A(e_M, m), C)$ is L. Let $w = w_1 w_2 \cdots w_n \in \Sigma^*$. There is a unique accepting path in $A(e_M, h_2(w))$ (and in no other $A(e_M, m)$) labeled w, and it is going successively through the states $(h_2(\varepsilon), h_2(w)) = (e_M, h_2(w))$, $(h_2(w_1), h_2(w_2 \cdots w_n))$, $\ldots$, $(h_2(w), e_M) = (h_2(w), h_2(\varepsilon))$. For this path, the sum computed by the semilinear set is:

$$\sum_{i=1}^{n} h_2(w_1 \cdots w_{i-1}) \cdot h_1(w_i) \cdot h_2(w_{i+1} \cdots w_n) \ ,$$

at the point (e_M, e_M). This is precisely $[h_1(w)](e_M, e_M)$, and checking whether it is in T amounts to checking whether $h(w) \in \mathcal{T}$, thus $L = \bigcup_{m \in M} L(A(e_M, m), C)$. $\square$

We derive an interesting property of the logical characterization and circuit complexity of UnCA. Let MSO[<] be the monadic second-order logic with < as the unique numerical predicate, and FO+G[<] be the first-order logic with group quantifiers and $<$ as the unique numerical predicate. Both logics express exactly the regular languages (these are respectively the classical results of Büchi [18] and Barrington, Immerman, Straubing [19]). Now define the *extended majority* quantifier $\widehat{\mathrm{Maj}}$, introduced in [20], as: $w \models \widehat{\mathrm{Maj}}\ x\ \langle \varphi_i \rangle_{i=1,\ldots,m}$ iff $\sum_{x=1}^{|w|} |\{i \mid w \models \varphi_i(x)\}| - |\{i \mid w \not\models \varphi_i(x)\}| > 0$. Then:

Corollary 1. *A language is in $\mathcal{L}_{\mathrm{UnCA}}$ iff it can be expressed as a Boolean combination of formulas of the form:*

$$\widehat{\mathrm{Maj}}\ x\ \langle \varphi_i \rangle_{i=1,\ldots,m}$$

where each φ_i is an MSO[<] *formula or an* FO+G[<] *formula. Hence, $\mathcal{L}_{\mathrm{UnCA}} \subsetneq \mathrm{NC}^1$.*

Proof. We first show that the languages recognized by $(\mathbb{Z}, \mathbb{Z}^+)$ are those expressible as a formula of the form (or negation of) $\widehat{\mathrm{Maj}}\ x\ \langle Q_{A_i} x \rangle_{i=1,\ldots,m}$ where $A_i \subseteq \Sigma$, and $Q_{A_i} x$ is short for $\bigvee_{a \in A_i} Q_a x$.

Let $L \in \mathcal{L}((\mathbb{Z}, \mathbb{Z}^+))$, i.e., let $h\colon \Sigma^* \to \mathbb{Z}$ be a morphism and suppose $L = h^{-1}(\mathbb{Z}^+)$ (if $L = h^{-1}(\mathbb{Z}_0^-)$, then the negation of the formula we obtain here will describe L). We suppose moreover, w.l.o.g., that each $h(a)$, $a \in \Sigma$, is even. Now let m be $\max\{|h(a)| \mid a \in \Sigma\}$ and define, for $1 \leq i \leq m$:

$$A_i = \{a \in \Sigma \mid m + h(a) \geq 2 \times i\} \ .$$

Now let $w \in \Sigma^*$ be a word and $1 \leq x \leq |w|$. Then it holds that:

$$h(w_x) = \underbrace{|\{i \mid w_x \in A_i\}|}_{(m+h(a))/2} - \underbrace{|\{i \mid w_x \notin A_i\}|}_{m-(m+h(a))/2} \ .$$

Thus for $w \in \Sigma^*$, $h(w) > 0$ iff $w \models \widehat{\mathrm{Maj}}\ x\ \langle Q_{A_i} x \rangle_{i=1,\ldots,m}$, thus the language expressed by this latter formula is $h^{-1}(\mathbb{Z}^+) = L$.

Conversely, consider a formula $\widehat{\mathrm{Maj}}\ x\ \langle Q_{A_i} x \rangle_{i=1,\ldots,m}$. Then let $h\colon \Sigma^* \to \mathbb{Z}$ be the morphism defined by $h(a) = |\{i \mid a \in A_i\}| - |\{i \mid a \notin A_i\}|$, for $a \in \Sigma$. We have that for $w \in \Sigma^*$, $h(w) > 0$ iff the formula under consideration holds true, implying that the language recognized by the formula is $h^{-1}(\mathbb{Z}^+)$.

It follows that the languages recognized by $\mathbf{Z}^+$ are the Boolean combinations of languages expressible as such formulas. Now the languages (with one free variable) recognized by finite monoids are those recognized by MSO[<] or FO+G[<] formulas. Thus the *block product principle* [21, Theorem 3.40] implies that the languages of $\mathcal{L}_{\mathrm{UnCA}} = \mathcal{L}(\mathbf{Z}^+ \Box \mathbf{M})$ are those expressible as Boolean combinations of formulas of the form of the statement of the lemma. Similarly, the regular languages (with one free variable) are recognized by NC^1 circuits, and a formula or negation of a formula of the form $\widehat{\mathrm{Maj}}\ x\ \langle Q_{A_i} x \rangle_{i=1,\ldots,m}$ can be expressed by a threshold circuit. Now [21, Lemma 4.29] implies that $\mathcal{L}_{\mathrm{UnCA}} \subseteq \mathrm{NC}^1$. Strictness is implied by Theorem 3. $\square$

3.3 Capturing DetAPA and UnAPA

Write $\mathfrak{Z}^+(k)$ for the type set of $(\mathbb{Z}, \mathbb{Z}^+)^k$, that is, the sets expressible as a Boolean combination of expressions of the form $x_i > 0$. Let $\mathbf{ZMat}^+$ be the set of typed monoids $\{(\mathcal{M}_\mathbb{Z}(k), \mathfrak{Z}^+(k \times k)) \mid k \geq 1\}$, then:

Theorem 5. $\mathcal{L}(\mathbf{ZMat}^+) = \mathcal{L}_{\text{DetAPA}}$.

Proof. ($\mathcal{L}_{\text{DetAPA}} \subseteq \mathcal{L}(\mathbf{ZMat}^+)$) This is a direct consequence of Lemma 3.

($\mathcal{L}(\mathbf{ZMat}^+) \subseteq \mathcal{L}_{\text{DetAPA}}$) Given $k \geq 1$, a type $\mathcal{Z}$ of $(\mathbb{Z}, \mathbb{Z}^+)^{k \times k}$, and a morphism $h\colon \Sigma^* \to \mathcal{M}_\mathbb{Z}(k)$, we build a two-state DetAPA of dimension k^2 for $h^{-1}(\mathcal{Z})$. First, let $h'\colon \Sigma^* \to \mathcal{M}_\mathbb{Z}(k^2)$ be such that $h'(a)$ is the Kronecker product of the identity matrix of dimension k and $h(a)$. Define $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k)$ where each $\mathbf{e}_i$ is of dimension k. Then for any word w, $h(w) \in \mathcal{Z}$ iff $h'(w).\mathbf{e} \in \mathcal{Z}$. Now let $A = (\{r, s\}, \Sigma, \delta, r, \{s\})$, with $\delta = \{r, s\} \times \Sigma \times \{s\}$. Then let $U\colon \delta^* \to \mathcal{F}_{k^2}$ for $q \in \{r, s\}$, $a \in \Sigma$, and $\mathbf{x} \in \mathbb{Z}^{k^2}$ be defined by:

$$U_{q \xleftarrow{} a \xrightarrow{} s}(\mathbf{x}) = \begin{cases} h'(a).\mathbf{e} & \text{if } q = r, \\ h'(a).\mathbf{x} & \text{otherwise.} \end{cases}$$

This implies that for $w \in \Sigma^+$ and π its unique accepting path in A, it holds that $U_\pi(\mathbf{0}) = h'(w).\mathbf{e}$. Thus $L(A, U, \mathcal{Z}) = h^{-1}(\mathcal{Z})$. $\square$

Theorem 6. $\mathcal{L}(\mathbf{ZMat}^+ \Box \mathbf{M}) = \mathcal{L}_{\text{UnAPA}}$.

Proof. $\mathcal{L}_{\text{UnAPA}} \subseteq \mathcal{L}(\mathbf{ZMat}^+ \Box \mathbf{M})$ is the same as $\mathcal{L}_{\text{UnCA}} \subseteq \mathcal{L}(\mathbf{Z}^+ \Box \mathbf{M})$ in Theorem 4, thanks to Lemma 3.

$\mathcal{L}(\mathbf{ZMat}^+ \Box \mathbf{M}) \subseteq \mathcal{L}_{\text{UnAPA}}$ is the same as $\mathcal{L}(\mathbf{Z}^+ \Box \mathbf{M}) \subseteq \mathcal{L}_{\text{UnCA}}$ in Theorem 4 for the automaton part, and the same as Theorem 5 for the constraint set and affine function parts. $\square$

Now, applying the same arguments as in [22, Lemma 5], we have that DetAPA can simulate unambiguity, and thus $\mathcal{L}_{\text{UnAPA}} = \mathcal{L}_{\text{DetAPA}}$. This translates nicely in the algebraic framework thanks to Theorem 1:

Theorem 7. *The smallest variety containing* $\mathbf{ZMat}^+ \Box \mathbf{M}$ *is equal to that containing* $\mathbf{ZMat}^+$.

4 Formal power series

In this section, we show that the languages of DetAPA are those expressible as a Boolean combination of positive supports of $\mathbb{Z}$-valued rational series. This helps us derive a separation over the unary languages between $\mathcal{L}_{\text{CA}}$ and $\mathcal{L}_{\text{DetAPA}}$ — the separation was known ([2, Proposition 28]), but not over unary languages.

Definition 3 (e.g., [23]). *Functions from* Σ^* *into* $\mathbb{Z}$ *are called (*$\mathbb{Z}$*)-*series. *For such a series* r*, it is customary to write* (r, w) *for* $r(w)$*. We write* $\mathsf{supp}_+(r)$ *for the* positive support *of* r*, i.e.,* $\{w \mid (r, w) > 0\}$.

A linear representation *of dimension $k \geq 1$ is a triple* $(\mathbf{s}, h, \mathbf{g})$ *such that* $\mathbf{s} \in \mathbb{Z}^k$ *is a* row *vector,* $\mathbf{g} \in \mathbb{Z}^k$ *is a* column *vector, and* $h\colon \Sigma^* \to (\mathbb{Z}^{k\times k}, .)$ *is a monoid morphism, where* . *is the* usual *matrix multiplication. It defines the series* $r = ||(\mathbf{s}, h, \mathbf{g})||$ *with* $(r, w) = \mathbf{s}.h(w).\mathbf{g}$.

A series is said to be rational *if it is defined by a linear representation. We write* $\mathbb{Z}^{\mathrm{rat}}\langle\!\langle\Sigma^*\rangle\!\rangle$ *for the set of rational series.*

For a class $\mathcal{C}$ of languages, write $\mathsf{BC}(\mathcal{C})$ for the Boolean closure of $\mathcal{C}$. Arguments similar to those used in proving Theorem 5 allow us to show:

Theorem 8. *Over any alphabet* Σ, $\mathcal{L}_{\mathrm{DetAPA}} = \mathsf{BC}(\mathsf{supp}_+(\mathbb{Z}^{\mathrm{rat}}\langle\!\langle\Sigma^*\rangle\!\rangle))$.

Proof. ($\mathcal{L}_{\mathrm{DetAPA}} \subseteq \mathsf{BC}(\mathsf{supp}_+(\mathbb{Z}^{\mathrm{rat}}\langle\!\langle\Sigma^*\rangle\!\rangle))$) First note that there is a rational series r such that $\mathsf{supp}_+(r) = \{\varepsilon\}$. Let L be in $\mathcal{L}_{\mathrm{DetAPA}}$; we may thus suppose that $\varepsilon \notin L$. By the same token as in the proof of Theorem 5, there is a morphism $h\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(k)$, for some k, a vector $\mathbf{v} \in \{0,1\}^k$, and a type $\mathcal{Z}$ of $(\mathbb{Z}, \mathbb{Z}^+)^k$ such that:

$$L = \{w \mid h(w).\mathbf{v} \in \mathcal{Z}\}\ .$$

Further, similar to Lemma 2, $L(A, U, C_1 \oslash C_2) = L(A, U, C_1) \oslash L(A, U, C_2)$, for $\oslash \in \{\cup, \cap\}$ and any DetAPA $(A, U, C_1 \oslash C_2)$. Moreover, $L(A, U, \overline{C}) = \overline{L(A, U, C)} \cap L(A)$. We may thus suppose that $\mathcal{Z}$ is reduced to $\mathbb{Z}^{i-1} \times \mathbb{Z}^+ \times \mathbb{Z}^{k-i}$ for some i.

Now let h' be the morphism from Σ^* to $(\mathbb{Z}^{k\times k}, .)$, with . the *usual* matrix multiplication, where $h'(a) = (h(a))^{\mathrm{T}}$, with $a \in \Sigma$ and M^{T} the transpose of M. Note that $h(a_1a_2) = h(a_2).h(a_1) = ((h(a_1))^{\mathrm{T}}.(h(a_2))^{\mathrm{T}})^{\mathrm{T}}$, which is $(h'(a_1a_2))^{\mathrm{T}}$; more generally, $h(w) = (h'(w))^{\mathrm{T}}$. Thus we have that $\mathbf{v}^{\mathrm{T}}.h'(w) = (h(w).\mathbf{v})^{\mathrm{T}}$. Hence with $\mathbf{s} = \mathbf{v}^{\mathrm{T}}$ and $\mathbf{g}$ the column vector $\mathbf{e}_i$, $\mathbf{s}.h'(w).\mathbf{g} > 0$ iff $h(w).\mathbf{v} \in \mathcal{Z}$.

Now the triple $(\mathbf{s}, h', \mathbf{g})$ is a linear representation of a rational series which associates w to $\mathbf{s}.h'(w).\mathbf{g}$, and this concludes the proof.

($\mathsf{BC}(\mathsf{supp}_+(\mathbb{Z}^{\mathrm{rat}}\langle\!\langle\Sigma^*\rangle\!\rangle)) \subseteq \mathcal{L}_{\mathrm{DetAPA}}$) As $\mathcal{L}_{\mathrm{DetAPA}}$ is closed under union, complement, and intersection, we need only show that $\mathsf{supp}_+(\mathbb{Z}^{\mathrm{rat}}\langle\!\langle\Sigma^*\rangle\!\rangle) \subseteq \mathcal{L}_{\mathrm{DetAPA}}$.

Let $(\mathbf{s}, h, \mathbf{g})$ be a linear representation of dimension k of a rational series r over the alphabet Σ. Define $h'\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(k)$ by letting $h'(a) = (h(a))^{\mathrm{T}}$, for $a \in \Sigma$. Then for $w \in \Sigma^*$, $h(w) = (h'(w))^{\mathrm{T}}$. Now the rest of the proof is similar to that of Theorem 5: define $A = (\{r, t\}, \Sigma, \delta, r, \{r, t\})$, with $\delta = \{r, t\} \times \Sigma \times \{t\}$. Then let $U\colon \delta^* \to \mathcal{F}_k$ for $q \in \{r, t\}$, $a \in \Sigma$, and $\mathbf{x} \in \mathbb{Z}^k$, be defined by:

$$U_{q \xrightarrow{a} t}(\mathbf{x}) = \begin{cases} h'(a).\mathbf{s} & \text{if } q = r, \\ h'(a).\mathbf{x} & \text{otherwise.} \end{cases}$$

This implies that for $w \in \Sigma^*$ and π its unique accepting path in A, it holds that $U_\pi(\mathbf{0}) = \mathbf{s}.h(w)$. Thus letting $C = \{\mathbf{x} \mid \mathbf{x}.\mathbf{g} > 0\}$, with $\mathbf{x}$ a row vector and $\mathbf{g}$ a column vector, we have that $L(A, U, C) = \mathsf{supp}_+(r)$. ☐

Remark 1. The class of positive supports of $\mathbb{Z}$-rational series is the class of $\mathbb{Q}$-*stochastic languages* (see, e.g., [24]). As we are interested in showing that

$\mathcal{L}_{\text{DetAPA}}$ is not closed under concatenation, it is worth noting that $\mathbb{Q}$-stochastic languages are not closed under concatenation. We mention three proofs of this fact. Two proofs [25,24] show that $\mathbb{Q}$-stochastic languages are not closed under concatenation with a *finite* language; such a concatenation is expressible as a finite union of $\mathbb{Q}$-stochastic languages, and is thus not directly applicable to our case. A third proof [26] shows that the $\mathbb{Q}$-stochastic language $L = \{a^i\#(a+\#)^*\#a^i \mid i \in \mathbb{N}\}$ is such that $L \cdot \{a, \#\}^*$ is not $\mathbb{Q}$-stochastic. We conjecture that $L \cdot \{a, \#\}^*$ is neither in $\mathcal{L}_{\text{DetAPA}}$, but the proof given in [26] does not apply directly to our case. Finally, we note that the fact that unary $\mathbb{Q}$-stochastic languages are not closed under union [24] implies, as any regular language is $\mathbb{Q}$-stochastic, that there are nonregular unary languages in $\mathcal{L}_{\text{DetAPA}}$. We give a simple proof of this latter fact:

Corollary 2. *There is a nonregular unary language in* $\mathcal{L}_{\text{DetAPA}}$.

Proof. For $\Sigma = \{a\}$, and a series r over Σ, write $c_n = (r, a^n)$.

Suppose the sign of c_n depends only on the sign of $\sin(nx)$ for some real number x. Suppose $\mathsf{supp}_+(r)$ is regular, there is a set $E = \{k_1 + t.k_2 \mid t \in \mathbb{N}\}$, with $k_1 \in \mathbb{N}$, $k_2 \in \mathbb{N}^+$, such that for any $e \in E$, $c_e > 0$. Now this means that $\sin(k_1x + t \times k_2x)$ should be positive for any $t \in \mathbb{N}$. Let $z \in \mathbb{N}$, then note that for any real v:

$$\sin(k_1x + \frac{v - k_1x + 2\pi z}{k_2x} \times k_2x) = \sin(v) \ \ .$$

Thus, for any $z \in \mathbb{N}$, any integer t in the range:

$$\left]\frac{-\pi - k_1x + 2\pi z}{k_2x}, \frac{-k_1x + 2\pi z}{k_2x}\right[\ \ ,$$

is such that $\sin(k_1x + t \times k_2x) < 0$.

Suppose moreover that x/π is irrational. We reach a contradiction by showing that one such range contains an integer. Indeed, write $A = \frac{2\pi}{k_2x}$, $B = \frac{-\pi-k_1x}{k_2x}$, and $C = \frac{\pi}{k_2x}$; we are searching for a $z \in \mathbb{N}$ such that there is an integer in $[z.A + B, z.A + B + C]$. Now, as A is irrational, for any real interval $[a, b] \subseteq [0, 1]$ there is an integer z such that the fractional part of $z.A$ is in $[a, b]$ (see, e.g., [27, Th. 3.2]). Thus we can find a z such that there is an integer t with the property that $z.A + B \in [t - C/2 - C/4, t - C/2 + C/4]$. This implies that $z.A + B < t < z.A + B + C$, proving the claim.

Thus if c_n depends only on the sign of $\sin(nx)$ for some irrational number x, then $\mathsf{supp}_+(r)$ is nonregular. We give an example of one such r. Recall that r is rational iff the sequence of c_n's satisfies a linear recurrence relation [23]. Let r be specified by the linear recurrence relation and initial values:

$$c_n = 2 \times c_{n-1} - 5 \times c_{n-2}, \quad c_0 = 0, c_1 = 1 \ \ .$$

Then $c_n = \frac{1}{4}i((1-2i)^n - (1+2i)^n) = \frac{1}{2}5^{n/2}\sin(n\tan^{-1}(2))$. Then $x = \tan^{-1}(2)$ is such that x/π is not a rational number: this is a consequence of Niven's theorem [28, Cor. 3.12], asserting that if x/π is rational, then the only rational values of $\tan(x)$ are 0 and ± 1. $\square$

Let $\#\mathrm{NC}^1$ be the class of functions computed by DLOGTIME-uniform arithmetic circuits of polynomial size and logarithmic depth and PNC^1 be the class of languages expressible as $\{w \mid f(w) > 0\}$ for $f \in \#\mathrm{NC}^1$ (see [11]). Note that this class is included in L. As iterated matrix multiplication can be done in $\#\mathrm{NC}^1$ and PNC^1 is closed under the Boolean operations, it is readily seen from Theorem 8 that:

Corollary 3. $\mathcal{L}_{\mathrm{DetAPA}} \subseteq \mathrm{PNC}^1$.

Conclusion

Connections between variants of the Parikh automaton and complexity classes were investigated. In particular, natural characterizations of the language classes defined by deterministic and unambiguous constrained automata, in the theory of typed monoids, were obtained. We hope that these characterizations will suggest refinements that may help to better understand classes such as PNC^1 and NC^1.

We note in conclusion that the unary languages in $\mathcal{L}_{\mathrm{DetAPA}}$, and indeed the bounded languages in $\mathcal{L}_{\mathrm{DetAPA}}$, can be shown to belong to the DLOGTIME-DCL-uniform variant of NC^1. Recall that the latter is not known to equal what is commonly referred to as DLOGTIME-uniform NC^1 (see [30, P. 162]), or ALOGTIME. Yet we were unable to show that the unary languages in $\mathcal{L}_{\mathrm{DetAPA}}$ belong to the latter. Do they?

Additional preliminaries and proofs

1 Preliminaries

We first present some key concepts of algebraic language theory.

Varieties We say that a monoid M divides a monoid N if M is the morphic image of a submonoid of N. A class of languages is said to be a *variety of languages* if it is closed under the Boolean operations, inverse morphisms, and quotient by a word. A class of monoids is said to be a *(pseudo)variety of monoids* if it is closed under division and direct products. Eilenberg's theorem [9] states that the varieties of regular languages and of finite monoids are in one-to-one correspondence.

Bilateral semidirect product, wreath and block products Let M and N be finite monoids. To distinguish the operation of M and N, we denote the operation of M as $+$ and its identity element as 0 (although this operation is not necessarily commutative) and the operation of N implicitly and its identity element as 1. A left action of N on M is a function mapping pairs $(n, m) \in N \times M$ to $nm \in M$ and satisfying $n(m_1 + m_2) = nm_1 + nm_2$, $n_1(n_2 m) = (n_1 n_2)m$, $n0 = 0$ and $1m = m$. Right actions are defined symmetrically. If we have both a right and a left action of N on M that further satisfy $n_1(mn_2) = (n_1 m)n_2$, we define the bilateral semidirect product $M ** N$ as the monoid with elements in $M \times N$ and multiplication defined as $(m_1, n_1)(m_2, n_2) = (m_1 n_2 + n_1 m_2, n_1 n_2)$. This operation is associative and $(0, 1)$ acts as an identity for it. Given only a left action, the unilateral semidirect product $M * N$ is the bilateral semidirect product $M ** N$ where the right action on M is trivial ($mn = m$).

Let M, N be two monoids. The *wreath product* of M and N, written $M \wr N$, is defined as the unilateral semidirect product of M^N and N, where the left action of N on M^N is given by $(n \cdot f)(n') = f(n'n)$, for $f \colon N \to M$ and $n, n' \in N$. The *block product* of M and N, written $M \square N$, is defined as the bilateral semidirect product of $M^{N\times N}$ and N, where the right (resp. left) action of N on $M^{N\times N}$ is given by $(f \cdot n)(n_1, n_2) = f(n_1, nn_2)$ (resp. $(n \cdot f)(n_1, n_2) = f(n_1 n, n_2)$), for $f \colon N \times N \to M$ and $n, n_1, n_2 \in N$. Note the following property of multiplication in the block product. Let M and N be monoids, and $(f_i, n_i) \in M \square N$, $i = 1, \ldots, k$:

$$\prod_{i=1}^{k}(f_i, n_i) = \left(\sum_{i=1}^{k} (n_1 \cdots n_{i-1}) \cdot f_i \cdot (n_{i+1} \cdots n_k), n_1 n_2 \cdots n_k \right) .$$

where the sum of two functions is their pointwise product over M. A similar property holds for the wreath product, without the right action by $(n_{i+1} \cdots n_k)$. Also, we define the wreath product (resp. block product) of two sets of monoids in the natural way.

We now turn to the theory of typed monoids.

Definition 4 (Typed morphism [10]). *Let $(S, \mathfrak{S}), (T, \mathfrak{T})$ be typed monoids. A* typed morphism $h\colon (S, \mathfrak{S}) \to (T, \mathfrak{T})$ *is a pair $(h_S, h_{\mathfrak{S}})$ with $h_S\colon S \to T$ a monoid morphism, $h_{\mathfrak{S}}\colon \mathfrak{S} \to \mathfrak{T}$ a morphism of Boolean algebras, and:*

$$\forall \mathcal{S} \in \mathfrak{S}, h_S(\mathcal{S}) = h_{\mathfrak{S}}(\mathcal{S}) \cap h_S(S) \ .$$

This compatibility condition allows us to omit the indices of the morphisms. The typed morphism is injective if both h_S and $h_{\mathfrak{S}}$ are.

Definition 5 (Typed submonoids, varieties [10]). *A typed monoid $(T, \mathfrak{T})$ is a* typed submonoid *of $(S, \mathfrak{S})$ if T is a submonoid of S and there is an injective morphism from $(T, \mathfrak{T})$ to $(S, \mathfrak{S})$. A typed monoid $(T, \mathfrak{T})$* divides *a typed monoid $(S, \mathfrak{S})$ if $(T, \mathfrak{T})$ is a morphic image of a typed submonoid of $(S, \mathfrak{S})$.*

A variety of typed monoids *is a class of typed monoids closed under division, direct product, and* shifting*, that is the operation mapping $(S, \mathfrak{S})$ to $(S, \{\lambda^{-1}\mathcal{S}\rho^{-1} \mid \mathcal{S} \in \mathfrak{S}\})$, for some $\lambda, \rho \in S$, where $\lambda^{-1}\mathcal{S}\rho^{-1} = \{s \mid \lambda s \rho \in \mathcal{S}\}$.*

The usual wreath product (resp. block product) of M and N, i.e., the unilateral (resp. bilateral) semidirect product of M^N (resp. $M^{N\times N}$) and N, results, in the infinite monoid case, in monoids with uncountably many elements, failing to fall within the definition of typed monoid. Thus these products are restricted, in [8], to *type-respecting* functions, that is, functions that only depend on the type of their arguments (multiplied by some constants). Here, we are not concerned with this technicality as all our monoids N will be finite. Hence we define:

Definition 6 (Block product [8]). *Let $(M, \mathfrak{M})$, $(N, \mathfrak{N})$ be typed monoids. The* block product *of $(M, \mathfrak{M})$ and $(N, \mathfrak{N})$, written $(M, \mathfrak{M})\Box(N, \mathfrak{N})$, is $(M\Box N, \mathfrak{S})$ with $\mathfrak{S} = \{\mathcal{S}_{\mathcal{M}} \mid \mathcal{M} \in \mathfrak{M}\}$ where:*

$$\mathcal{S}_{\mathcal{M}} = \{(f, n) \in S \mid f(e_N, e_N) \in \mathcal{M}\} \ .$$

The primary focus of the theory of typed monoid was the algebraic characterization of classes of circuit complexity, which are naturally closed under reversal. We need a finer notion of product to characterize classes which are not closed under reversal. In the classical untyped theory, the appropriate algebraic construction for such classes relies on the *wreath product*, and we naturally derive the following definition in the typed case:

Definition 7 (Wreath product). *Let $(M, \mathfrak{M})$, $(N, \mathfrak{N})$ be typed monoids. The* wreath product *of $(M, \mathfrak{M})$ and $(N, \mathfrak{N})$, written $(M, \mathfrak{M}) \wr (N, \mathfrak{N})$, is $(M \wr N, \mathfrak{S})$ with $\mathfrak{S} = \{\mathcal{S}_{\mathcal{M}} \mid \mathcal{M} \in \mathfrak{M}\}$ where:*

$$\mathcal{S}_{\mathcal{M}} = \{(f, n) \in S \mid f(e_N) \in \mathcal{M}\} \ .$$

2 Normal forms of CA and APA

Lemma 4. *If $C \subseteq \mathbb{Z}^d$ is modulo-free, then for all $1 \le d' < d$ there is a basic set $B \subseteq \mathbb{Z}^d$ such that for all $\mathbf{x} \in \mathbb{Z}^{d'}$ and $1 \le i \le d - d'$:*

$$(\mathbf{x}, \mathbf{e}_i) \in B \Leftrightarrow (\mathbf{x}, \mathbf{e}_i) \in C \ .$$

Proof. Let C be expressed as a Boolean combination of expressions of the form $\sum_i \alpha_i x_i > c$. Let the negations be pushed to the lowest level of the Boolean combination, now:

$$\neg\left(\sum_t \alpha_t x_t > c\right) \sim \sum_i \alpha_i x_i < c+1 \sim \sum_i (-\alpha_i) x_i > -(c+1) \enspace ,$$

thus we may suppose that C is a *positive* Boolean combination. Now replace each expression $\sum_i \alpha_i x_i > c$ by $\sum_i \alpha'_i x_i > c$ where $\alpha'_i = \alpha_i$ if $i \leq d'$ and $\alpha'_i = \alpha_i - c$ otherwise. Thus let $\mathbf{x} \in \mathbb{Z}^{d'} \times \{\mathbf{e}_i \mid 1 \leq i \leq d - d'\}$. It holds that $\sum_i \alpha_i x_i > c$ iff $\sum_i \alpha'_i x_i > 0$, thus the set resulting from the replacement is a basic set with the property of the statement of the lemma. $\square$

Lemma 4. *Every DetCA (resp. UnCA) has the same language $L \subseteq \Sigma^+$ as another DetCA (resp. UnCA) (A, C) with $L(A) = \Sigma^*$ and C a basic set.*

Proof. Let (A, C) be a CA. Let C be expressed as a Boolean combination of expressions that use the relation symbols $<$, $\equiv_p$, function symbol $+$, and constants from $\mathbb{N}$. We first remove the $\equiv_p$ relation symbols. Consider an expression in the definition of C of the form $\sum_t \alpha_t \times x_t \equiv_p c$, for $\alpha_t \in \mathbb{Z}$, where the sum ranges over all transitions. Then define A' to be p copies $A_0, A_1, \ldots, A_{p-1}$ of A, the initial state being that of A_0, and the final states being the final states of all the A_i's. On taking a transition t of A in the copy A_i, the automaton A' jumps to the copy $A_{i+\alpha_t \pmod p}$. Thus A' ends its computation in the copy A_c iff the numbers x_t of times each transition t of A has been taken in any copy are such that $\sum_t \alpha_t \times x_t \equiv_p c$. Now, given $\mathsf{Pkh}(\pi)$ for a path π, a basic set can check in which state π ended (this is the only state entered more often than exited, see, e.g., [3, Claim 1]), thus the expression under consideration can be replaced by an expression checking whether the path ended in A_c. The other expressions are simply adjusted to be oblivious to the copies. By induction, we may suppose that C is thus defined without the relations $\equiv_p$, i.e., that C is modulo-free.

Thus let (A, C) be a CA with C modulo-free. We rely on Lemma 4 to turn C into a basic set. Write $A = (Q, \Sigma, \delta, q_0, F)$, then define $A' = (Q \cup \{s\}, \Sigma, \delta', s, F)$ where $s \notin Q$, $\delta' = \delta \cup T$ with $T = \{s \bullet\!\!- a \!\rightarrow q \mid q_0 \bullet\!\!- a \!\rightarrow q \in \delta\}$. Order δ' so that the transitions of T appear last. Now define C' as being C where each expression $\sum_{t \in \delta} \alpha_t x_t > c$ is replaced by $\sum_{t \in \delta'} \alpha_t x_t > 0$ where α_t for $t = s \bullet\!\!- a \!\rightarrow q \in T$ is $\alpha_{q_0 \bullet\!- a \rightarrow q}$. Clearly $L(A, C) = L(A', C')$. Now exactly one of the transitions in T will be taken in a nonempty run in A', thus $L(A', C') = L(A, C' \cap \{\mathbf{e}_i \mid 1 \leq i \leq |T|\})$, and by Lemma 4, there is a basic set B such that $L(A', C') = L(A', B)$.

Note that all the operations made on the automaton part of the CA do not change whether it is deterministic or unambiguous. We now make sure that the language of the underlying automaton is Σ^*.

Let (A, C) be a DetCA with C basic. Following Karianto [4], we can set all the states of A to be final if we tweak C to check that the ending state of the path is final in A. This is equivalent to checking which state was entered more often than exited, which is expressed, for a state q different from the initial state as $\sum_{t\in q^-} x_t - \sum_{t\in q^+} x_t > 0$, where q^- (resp. q^+) is the set of transitions going to (resp. leaving) the state q. Thus the set K of Parikh images of accepting paths in A is basic, and we can replace C by $C \cap K$, a basic set.

For the UnCA case, this is shown in [3, Proposition 1]. Note that the modification made to C (which is essentially to replace it with $C \times \mathbb{Z}^+$) preserves the fact that the constraint set is basic. $\square$

Lemma 4. *For $(A, C_1 \cap C_2)$ a DetCA or an UnCA it holds that:*

$$L(A, C_1 \cap C_2) = L(A, C_1) \cap L(A, C_2) \enspace .$$

The same holds for $\cup$.

Proof. For $\cup$, this is true of DetCA, UnCA, and CA. For $\cap$, the inclusion from right to left holds for the three models; the converse inclusion only holds for DetCA and UnCA. Let $w \in L(A, C_1) \cap L(A, C_2)$ for A a deterministic or unambiguous automaton. Then there is only one accepting path π in A with label w, and thus $\mathsf{Pkh}(\pi) \in C_1 \cap C_2$, and in turn $w \in L(A, C_1 \cap C_2)$. $\square$

Lemma 4. *Let $L \subseteq \Sigma^+$ be in $\mathcal{L}_{\mathrm{DetAPA}}$. There is a morphism $h\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(k)$, for some k, and a set $\mathcal{Z} \subseteq \mathbb{Z}^{k^2}$ expressible as a Boolean combination of expressions $x_i > 0$, such that $L = h^{-1}(\mathcal{Z})$.*

Similarly, let $L \subseteq \Sigma^+$ be in $\mathcal{L}_{\mathrm{APA}}$ (resp. in $\mathcal{L}_{\mathrm{UnAPA}}$). There is an automaton (resp. unambiguous automaton) A with transition set δ, a morphism $h\colon \delta^* \to \mathcal{M}_{\mathbb{Z}}(k)$, for some k, and a set $\mathcal{Z} \subseteq \mathbb{Z}^{k^2}$ expressible as a Boolean combination of expressions $x_i > 0$, such that $L = \mu_A(h^{-1}(\mathcal{Z}) \cap \mathsf{Run}(A))$.

Proof. Using [2, Lemma 23], then [2, Remark 35], we obtain that for any $L \subseteq \Sigma^+$ in $\mathcal{L}_{\mathrm{DetAPA}}$, there is a morphism $g\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(n)$, for some n, a vector $\mathbf{s} \in \mathbb{Z}^k$, and a modulo-free set C such that:

$$L = \{w \mid g(w).\mathbf{s} \in C\} \enspace .$$

Similarly, using [2, Lemma 24], then [2, Lemma 23], we obtain that for any $L \subseteq \Sigma^+$ in $\mathcal{L}_{\mathrm{APA}}$ (resp. $\mathcal{L}_{\mathrm{UnAPA}}$), there is an automaton (resp. unambiguous automaton) A with transition set δ, a morphism $g\colon \delta^* \to \mathcal{M}_{\mathbb{Z}}(n)$, for some n, a vector $\mathbf{s} \in \mathbb{Z}^k$, and a modulo-free set C such that:

$$L = \mu_A(\{\pi \in \mathsf{Run}(A) \mid g(\pi).\mathbf{s} \in C\}) \enspace .$$

The rest of this proof will focus on reaching the statement of the lemma for DetAPA; the proof is identical for UnAPA and APA, as A is left unchanged.

We first show that we can turn C into a basic set using Lemma 4. Extend g to a morphism $g'\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(n+1)$ defined by:

$$g'(a) = \begin{pmatrix} \boxed{g(a)} & \begin{matrix} 0 \\ \vdots \end{matrix} \\ 0 \cdots & 1 \end{pmatrix} \enspace ,$$

and let $\mathbf{s}'$ be the vector $(\mathbf{s}, 1)$. Then for any nonempty path π in A, $g'(\pi).\mathbf{s}' = (g(\pi).\mathbf{s}, 1)$. Moreover, by Lemma 4, there is a basic set B such that $B \cap \mathbb{Z}^n \times \{1\} = C \times \{1\}$. Thus $g(\pi).\mathbf{s} \in C$ iff $g'(\pi).\mathbf{s}' \in B$.

Thus we suppose that C is a basic set. We show that we can turn C into a set of dimension n' expressible as a Boolean combination of expressions $x_i > 0$. Suppose $\sum_i \alpha_i x_i > 0$ appears in the expression defining C. We extend g so that this sum is computed by the matrices. For $a \in \Sigma$, write the lines of $g(a)$ as $L_1, L_2, \ldots, L_n$. Then define $g' \colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(n+1)$ as the morphism mapping $a \in \Sigma$ to the matrix consisting of lines $(L_1, 0), (L_2, 0), \ldots, (L_n, 0), (\sum_{i \leq n} \alpha_i L_i, 0)$. This is such that, for $w \in \Sigma^*$, the last component of $g(w).(\mathbf{s}, 0)$ is indeed the sum under consideration. Thus the formula $\sum_i \alpha_i x_i > 0$ can be replaced by $x_{n+1} > 0$. Proceeding inductively results in a set $C \subseteq \mathbb{Z}^{n'}$ expressible as a Boolean combination of expressions of the form $x_i > 0$.

Thus we suppose that C is expressible as a Boolean combination of expressions $x_i > 0$. We show that the product $g(w).\mathbf{s}$ can be computed within the matrices. For $a \in \Sigma$, write the lines of $g(a)$ as $L_1, L_2, \ldots, L_n$, and define $h \colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(n+1)$ as the morphism mapping $a \in \Sigma$ to the matrix consisting of lines $(L_1, L_1.\mathbf{s}), (L_2, L_2.\mathbf{s}), \ldots, (L_n, L_n.\mathbf{s}), \mathbf{0}$. Then for $w \in \Sigma^+$, we have that $h(w).\mathbf{e}_{n+1} = (g(w).\mathbf{s}, 0)$. Thus let $\mathcal{Z} \subseteq \mathbb{Z}^{(n+1)\times(n+1)}$ be defined as $\mathcal{Z} = \mathbb{Z}^{n\times(n+1)} \times C \times \mathbb{Z}$. Then $g(w).\mathbf{s} \in C$ iff $h(w) \in \mathcal{Z}$, proving the lemma. $\square$

3 Capturing Parikh automata classes algebraically

Theorem 8. *There is a language $L \notin \mathcal{L}_{\mathrm{CA}}$ recognized by*[4] *$U_1 \wr (\mathbb{Z}, \mathbb{Z}^+)$ and by $(\mathbb{Z}, \mathbb{Z}^+) \wr (\mathbb{Z}, \mathbb{Z}^+)$.*

Proof. We treat the case $(\mathbb{Z}, \mathbb{Z}^+) \wr (\mathbb{Z}, \mathbb{Z}^+)$, this is similar for $U_1 \wr (\mathbb{Z}, \mathbb{Z}^+)$. We consider the language:

$$P_1 = \{w = w_1 w_2 \cdots w_k \in \{\sqsubset, \sqsupset\}^* \mid (\forall i)[|w_1 w_2 \cdots w_i|_\sqsubset \geq |w_1 w_2 \cdots w_i|_\sqsupset]\} ,$$

which is not in $\mathcal{L}_{\mathrm{CA}}$ [3, Proposition 5]. Define $h \colon \{\sqsubset, \sqsupset\}^* \to \mathbb{Z} \wr \mathbb{Z}$ by:

$$h(\sqsubset) = (f, -1)$$
$$h(\sqsupset) = (f, 1)$$
$$f(n) = \begin{cases} 0 & \text{if } n \leq 0 \ , \\ 1 & \text{otherwise.} \end{cases}$$

Note that f is indeed type-respecting in the sense of [8], as its value depends only on whether its argument is in the type $\mathbb{Z}^+$ of the typed monoid $(\mathbb{Z}, \mathbb{Z}^+)$. Now $h(w) = (f, |w|_\sqsupset - |w|_\sqsubset)$ and $f(0)$ is zero or less iff for all i it holds that $|w_1 w_2 \cdots w_i|_\sqsubset \geq |w_1 w_2 \cdots w_i|_\sqsupset$. Thus:

$$P_1 = h^{-1}(\{(f, n) \mid f(0) \in \mathbb{Z}_0^-\}) ,$$

[4] The wreath product of two infinite monoids is ill-defined; however Theorem 3 stays true with an adequate definition mimicking that of [8].

and hence $P_1 \in \mathcal{L}((\mathbb{Z}, \mathbb{Z}^+) \wr (\mathbb{Z}, \mathbb{Z}^+))$. $\square$

Remark 2. The characterization of $\mathcal{L}_{\mathrm{UnCA}}$ by means of block products allows for an alternative algebra-based proof that $\mathcal{L}_{\mathrm{UnCA}}$ is closed under reversal (the proof of [3, Proposition 3], is automata-based). Let $L \in \mathcal{L}_{\mathrm{UnCA}}$, there is a finite monoid M, an integer k, a type $\mathcal{T}$ of $(\mathbb{Z}, \mathbb{Z}^+)^k \Box M$, and a morphism $h\colon \Sigma^* \to \mathbb{Z}^k \Box M$ such that $L = h^{-1}(\mathcal{T})$. Define M^{c} as M where the operation is reversed, and $h^{\mathrm{c}}\colon \Sigma^* \to \mathbb{Z}^k \Box M^{\mathrm{c}}$ by:

$$\begin{aligned} h^{\mathrm{c}}(a) &= (f_a^{\mathrm{c}}, \Psi_2(h(a))) \text{ where} \\ f_a^{\mathrm{c}}(m_1, m_2) &= [\Psi_1(h(a))](m_2, m_1) \ . \end{aligned}$$

Then the reversal of L is $(h^{\mathrm{c}})^{-1}(\mathcal{T})$ and, as M^{c} is also a finite monoid, it is in $\mathcal{L}(\mathbf{Z}^+ \Box \mathbf{M}) = \mathcal{L}_{\mathrm{UnCA}}$.

Similarly:

Corollary 4. $\mathcal{L}_{\mathrm{DetAPA}}$ *is closed under reversal.*

Proof. For a matrix M, write M^{T} for the transpose of M. We extend this notation to types in $\mathfrak{Z}^+(k)$ naturally. Let $L \in \mathcal{L}(\mathbf{ZMat}^+)$, there are $h\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(k)$ and $\mathcal{Z} \in \mathfrak{Z}^+(k \times k)$ such that $L = h^{-1}(\mathcal{Z})$. Define $h'\colon \Sigma^* \to \mathcal{M}_{\mathbb{Z}}(k)$ by $h'(a) = (h(a))^{\mathrm{T}}$. Then for a word w, $h(w) = (h'(w^{\mathrm{R}}))^{\mathrm{T}}$, and thus $h'(w) \in \mathcal{Z}^{\mathrm{T}}$ iff $h(w^{\mathrm{R}}) \in \mathcal{Z}$, where w^{R} is the reversal of w. Hence the reversal of L is $(h')^{-1}(\mathcal{Z}^{\mathrm{T}}) \in \mathcal{L}(\mathbf{ZMat}^+)$. $\square$

References

1. F. Klaedtke and H. Rueß, "Monadic second-order logics with cardinalities," in *30th Int. Coll. Automata, Languages and Programming*, 2003, pp. 681–696.
2. M. Cadilhac, A. Finkel, and P. McKenzie, "Affine parikh automata," *RAIRO - Theoretical Informatics and Applications*, vol. 46, no. 04, pp. 511–545, 2012. [Online]. Available: http://dx.doi.org/10.1051/ita/2012013
3. ——, "Unambiguous constrained automata," in *Developments in Language Theory*, ser. Lecture Notes in Computer Science, vol. 7410. Springer Berlin / Heidelberg, 2012, pp. 239–250. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-31653-1_22
4. W. Karianto, "Parikh automata with pushdown stack. Diploma thesis, RWTH Aachen," 2004.
5. D. A. M. Barrington, "Bounded-width polynomial size branching programs recognize exactly those languages in NC^1," *Journal of Computer and System Sciences*, vol. 38, pp. 150–164, 1989.
6. D. A. M. Barrington and D. Thérien, "Finite monoids and the fine structure of NC^1," *Journal of the Association of Computing Machinery*, vol. 35, pp. 941–952, 1988.
7. H. Straubing, *Finite Automata, Formal Logic, and Circuit Complexity*. Boston: Birkhäuser, 1994.
8. A. Krebs, K.-J. Lange, and S. Reifferscheid, "Characterizing TC^0 in terms of infinite groups," *Theory of Computing Systems*, vol. 40, no. 4, pp. 303–325, 2007. [Online]. Available: http://dx.doi.org/10.1007/s00224-006-1310-2
9. S. Eilenberg, *Automata, Languages, and Machines, Volume B*, ser. Pure and Applied Mathematics. Academic Press, 1976.
10. C. Behle, A. Krebs, and S. Reifferscheid, "Typed monoids - an Eilenberg-like theorem for non regular languages," in *Algebraic Informatics*, ser. Lecture Notes in Computer Science, vol. 6742. Springer, 2011, pp. 97–114.
11. H. Caussinus, P. McKenzie, D. Thérien, and H. Vollmer, "Nondeterministic NC^1 computation," *Journal of Computer and System Sciences*, vol. 57, pp. 200–212, 1998.
12. J. Gill, "Computational complexity of probabilistic turing machines," *SIAM Journal on Computing*, vol. 6, pp. 675–695, 1977.
13. N. Limaye, M. Mahajan, and B. V. R. Rao, "Arithmetizing classes around NC^1 and L," *Theory Comput. Syst.*, vol. 46, no. 3, pp. 499–522, 2010.
14. E. Allender, V. Arvind, and M. Mahajan, "Arithmetic complexity, Kleene closure, and formal power series," *Theory Comput. Syst.*, vol. 36, no. 4, pp. 303–328, 2003.
15. S. Ginsburg and E. H. Spanier, "Semigroups, Presburger formulas and languages," *Pacific Journal of Mathematics*, vol. 16, no. 2, pp. 285–296, 1966.
16. R. J. Parikh, "On context-free languages," *Journal of the ACM*, vol. 13, no. 4, pp. 570–581, 1966.
17. H. B. Enderton, *A Mathematical Introduction to Logic*. Academic Press, 1972.
18. J. R. Büchi, "Weak second order arithmetic and finite automata," *Z. Math. Logik Grundlagen Math.*, vol. 6, pp. 66–92, 1960.
19. D. A. M. Barrington, N. Immerman, and H. Straubing, "On uniformity within NC^1," *J. Comput. Syst. Sci*, vol. 41, no. 3, pp. 274–306, 1990. [Online]. Available: http://dx.doi.org/10.1016/0022-0000(90)90022-D
20. C. Behle, A. Krebs, and M. Mercer, "Linear circuits, two-variable logic and weakly blocked monoids," in *Mathematical Foundations of Computer Science*, ser. Lecture Notes in Computer Science, vol. 4708. Springer-Verlag, 2007, pp. 147–158.

21. A. Krebs, "Typed semigroups, majority logic, and threshold circuits," Ph.D. dissertation, Eberhard Karls University of Tübingen, 2008.
22. M. Cadilhac, A. Finkel, and P. McKenzie, "Bounded parikh automata," in *Proc. 8th Internat. Conference on Words*, ser. EPTCS, vol. 63, 2011, pp. 93–102.
23. J. Berstel and C. Reutenauer, *Noncommutative Rational Series with Applications*, ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010.
24. A. Salomaa and M. Soittola, *Automata-Theoretic Aspects of Formal Power Series*. New York: Springer, 1978.
25. M. Fliess, "Propriétés booléennes des langages stochastiques," *Mathematical Systems Theory*, vol. 7, no. 4, pp. 353–359, 1974. [Online]. Available: http://dx.doi.org/10.1007/BF01890611
26. P. Turakainen, "Some closure properties of the family of stochastic languages," *Information and Control*, vol. 18, no. 3, pp. 253–256, 1971.
27. I. Niven, *Diophantine Approximations*. Dover Publications, 1963.
28. ——, *Irrational Numbers*, ser. Carus Mathematical Monographs. Mathematical Association of America, 1956. [Online]. Available: http://books.google.ca/books?id=ov-IlIEo47cC
29. P. Flajolet and J.-M. Steyaert, "Complexity classes of languages and operators," IRIA, Tech. Rep. 92, 1974.
30. H. Vollmer, *Introduction to Circuit Complexity – A Uniform Approach*, ser. Texts in Theoretical Computer Science. Springer Verlag, 1999.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de