First-Order logic and its Infinitary Quantifier Extensions over Countable Words

Bharat Adsul*, Saptarshi Sarkar*, and A.V. Sreejith†

*IIT Bombay, India and †IIT Goa, India

Abstract. We contribute to the refined understanding of the language-logic-algebra interplay in the context of *first-order* properties of countable words. We establish decidable algebraic characterizations of one variable fragment of FO as well as boolean closure of existential fragment of FO via a strengthening of Simon's theorem about piecewise testable languages. We propose a new extension of FO which admits infinitary quantifiers to reason about the inherent infinitary properties of countable words. We provide a very natural and hierarchical block-product based characterization of the new extension. We also explicate its role in view of other natural and classical logical systems such as WMSO and FO[cut] - an extension of FO where quantification over Dedekind-cuts is allowed. We also rule out the possibility of a finite-basis for a block-product based characterization of these logical systems. Finally, we report simple but novel algebraic characterizations of one variable fragments of the hierarchies of the new proposed extension of FO.

Keywords: Countable words · First-order logic · Monoids.

1 Introduction

Over finite words, we have a foundational language-logic-algebra connection (see [18,10]) which equates regular-expressions, MSO-logic, and (recognition by) finite monoids/automata. In fact, one can effectively associate, to a regular language, its finite *syntactic monoid*. This canonical algebraic structure carries a rich amount of information about the corresponding language. Its role is highlighted by the classical Schutzenberger-McNaughton-Papert theorem (see, for instance, [11]) which shows that *aperiodicity* property of the syntactic monoid coincides with describability using star-free expressions as well as definability in First-Order (FO) logic. So, we arrive at a refined understanding of the language-logic-algebra connection to an important subclass of regular languages: it equates star-free regular expressions, FO-logic, and aperiodic finite monoids.

A variety of algebraic tools have been developed and crucially used to obtain deeper insights. Some of these tools [11,15,17] are: ordered monoids, the so-called Green's relations, wreath/block products and related principles etc. Let us mention Simon's celebrated theorem [14] - which equates piecewise-testable languages, Boolean closure of the existential fragment of FO-logic and J-trivial finite monoids[1]. It is important to note that this is an effective characterization,

[1] It refers to J - one of the fundamental Green's relations

that is, they provide a decidable characterization of the logical fragment. There have been several results of this kind (see the survey [6]). Another particularly interesting set of results is in the spirit of the fundamental Krohn-Rhodes theorem. These results establish a block-product based decompositional characterization of a logical fragment and have many important applications [15]. The prominent examples are a characterization of FO-logic (resp. FO^2, the two-variable fragment) in terms of strongly (resp. weakly) iterated block-products of copies of the unique 2-element aperiodic monoid.

One of the motivations for this work is to establish similar results in the theory of regular languages of countable words. We use the overarching algebraic framework developed in the seminal work [4] to reason about languages of countable words. This framework extends the language-logic-algebra interplay to the setting of countable words. It develops fundamental algebraic structures such as *finite* $\circledast$-monoids and $\circledast$-algebras and equates MSO-definability with recognizability by these algebraic structures. A detailed study of a variety of sub-logics of MSO over countable words is carried out in [5]. This study also extends classical Green's relations to $\circledast$-algebras and makes heavy use of it. Of particular interest to us are the results about algebraic equational characterizations of FO, FO[cut] – an *extension* of FO that allows quantification over Dedekind cuts and WMSO – an *extension* of FO that allows quantification over finite sets. A decidable algebraic characterization of FO^2 over countable words is also presented in [9]. Another recent development [1] is the seamless integration of block products into the countable setting. The work introduces the block product operation of the relevant algebraic structures and establishes an appealing block product principle. Further, it naturally extends the above-mentioned block product characterizations of FO and FO^2 to countable words.

In this work, we begin our explorations into the *small* fragments of FO over countable words, guided by the choice of results in [6]. We arrive at the language-logic-algebra connection for FO^1 – the one variable fragment of FO. Coupled with earlier results about FO^2 and $\mathrm{FO}{=}\mathrm{FO}^3$ (see [7]), this completes our algebraic understanding of FO fragments defined by the number of permissible variables. We next extend Simon's theorem on piecewise testable languages to countable words and provide a natural algebraic characterization of the Boolean closure of the existential-fragment of FO. Fortunately or unfortunately, depending on the point of view, this landscape of small fragments of FO over countable words parallels very closely the same landscape over finite words. This can be attributed to the limited expressive power of FO over countable words. For instance, Bès and Carton [3] showed that the seemingly natural 'finiteness' property (that the set of all positions is a finite set) of countable words can not be expressed in FO!

One of the main contributions of this work is the introduction of new *infinitary* quantifiers to FO. The works [2,8] also extend FO over arbitrary structures by *cardinality/finitary-counting* quantifiers and studies decidable theories thereof. An extension of FO over finite and ω-words by *modulus-counting quantifiers* is algebraically characterized in [16]. The main purpose of our new quantifiers is to naturally allow expression of infinitary features which are inherent in the countable setting and study the resulting definable formal languages in the algebraic framework of [4]. An example formula using such an infinitary quantifier is: $\exists^{\infty_1} x : a(x) \wedge \neg\exists^{\infty_1} x : b(x)$. In its natural semantics, this formula with one variable asserts that there are infinitely many a-labelled positions and only finitely many b-labelled positions. We propose an extension of FO called FO[∞] that supports *first-order* infinitary quantifiers of the form $\exists^{\infty_k} x$ to talk about existence of higher-level infinitely (more accurately, Infinitary rank k) many witnesses x. We organize FO[∞] in a natural hierarchy based on the maximum allowed infinitary-level of the quantifiers.

We now summarize the key technical results of this paper. We establish a hierarchical block product based characterization of FO[∞]. Towards this, we identify an appropriate simple family of $\circledast$-algebras and show that this family (in fact, its initial fragments) serve as a basis in our hierarchical block product based characterization. We establish that FO[∞] properties can be expressed simultaneously in FO[cut] as well as WMSO. We also show that the language-logic-algebra connection for FO^1 admits novel generalizations to the one variable fragments of the new extension of FO. We finally present 'no finite block product basis' theorems for our FO extensions, FO[cut], and the class FO[cut] $\cap$ WMSO. This is in contrast with [1] where the unique 2-element $\circledast$-algebra is a basis for a block-product based characterization of FO.

The rest of the paper is organized as follows. Section 2 recalls basic notions about countable words and summarizes the necessary algebraic background from the framework [4]. Section 3 deals with the small fragments of FO: FO^1 and the Boolean closure of the existential fragment of FO. Section 4 contains the extensions FO[∞] and results relevant to it. Section 5 is concerned with 'no finite block product basis' theorems.

2 Preliminaries

In this section we briefly recall the algebraic framework developed in [4].

Countable words A *countable linear ordering* (or simply ordering) $\alpha = (X, <)$ is a non-empty countable set X equipped with a total order: X is the *domain* of α. An ordering $\beta = (Y, <)$ is called a *subordering* of α if $Y \subseteq X$ and the order on Y is induced from that of X. We denote by $\omega, \omega^*, \delta, \eta$ the orderings $(\mathbb{N}, <), (-\mathbb{N}, <), (\mathbb{Z}, <), (\mathbb{Q}, <)$ respectively. A *Dedekind cut* (or simply a cut) is a left-closed subset $Y \subseteq X$ of α. Given disjoint linear orderings $(\beta_i)_{i\in\alpha}$ indexed with a linear ordering α, their *generalized sum* $\sum_{i\in\alpha}\beta_i$ is the linear ordering over the union of the domains of the β_i's, with the order defined by $x < y$ if either $x \in \beta_i$ and $y \in \beta_j$ with $i < j$, or $x, y \in \beta_i$ for some i, and $x < y$ in β_i. The book [12] contains a detailed study of linear orderings.

An *alphabet* Σ is a finite set of symbols called *letters*. Given a linear ordering α, a *countable word* (henceforth called word) over Σ of domain α is a mapping $w : \alpha \to \Sigma$. The domain of a word is denoted $dom(w)$. For a subset $I \subseteq dom(w)$, $w|_I$ denotes the *subword* got by restricting w to the domain I. If I is an *interval*

$(\forall x, y \in I,\ x < z < y \to z \in I)$ then $w|_I$ is called a *factor* of w. The set of all words is denoted $\Sigma^{\circledast}$ and the set of all non-empty (resp. finite) words $\Sigma^{\oplus}$ (resp. Σ^*). A *language* (of countable words) is a subset of $\Sigma^{\circledast}$. The *generalized concatenation* of the words $(w_i)_{i\in\alpha}$ indexed by a linear ordering α is $\prod_{i\in\alpha} w_i$ and denotes the word w of domain $\sum_{i\in\alpha} \beta_i$ where β_i are disjoint and such that $w|_{\beta_i}$ is isomorphic to w_i for all $i \in \alpha$.

The *empty word* ε, is the only word of empty domain. The *omega power* of a word w is defined as $w^{\omega} ::= \prod_{i\in\omega} w$. The *omega* power* of a word w, denoted by w^{ω^*}, is the concatenation of omega* many w's. The *perfect shuffle* for a non-empty finite set of letters $A \subseteq \Sigma$ (denoted by A^{η}) is a word of domain $(\mathbb{Q}, <)$ in which only letters from A occur and, all non-empty and non-singleton intervals contain at least one occurrence of each letter in A. This word is unique up to isomorphism [13]. We can extend the notion of perfect shuffle to a finite set of words $W = \{w_1, \ldots, w_k\}$. We define W^{η} to be $\prod_{i\in\mathbb{Q}} w_{f(i)}$ where $f : (\mathbb{Q}, <) \to \{1, 2, \ldots, k\}$ is the unique perfect shuffle over the set of letters $\{1, 2, \ldots, k\}$.

The algebra A *$\circledast$-monoid* $\mathbf{M} = (M, \pi)$ is a set M equipped with an operation π, called the *product*, from $M^{\circledast}$ to M, that satisfies $\pi(a) = a$ for all $a \in M$, and the *generalized associativity* property: for every words u_i over M with i ranging over a countable linear ordering α, $\pi\left(\prod_{i\in\alpha} u_i\right) = \pi\left(\prod_{i\in\alpha} \pi(u_i)\right)$. We reserve the notation $\mathtt{id}$ for the *identity element* $\mathtt{id} = \pi(\varepsilon)$; it is called the neutral element in [4]. An example of a $\circledast$-monoid is the free $\circledast$-monoid $(\Sigma^{\circledast}, \varepsilon, \prod)$ over the alphabet Σ with the product being the generalized concatenation. Now we discuss some natural algebraic notions. A *morphism* from a $\circledast$-monoid (M, π) to a $\circledast$-monoid (M', π') is a map $h : M \to M'$ such that, for every $w \in M^{\circledast}$, $h(\pi(w)) = \pi'(\bar{h}(w))$ where $\bar{h}$ is the pointwise extension of h to words. We skip the notions sub-*$\circledast$-monoid* and direct products since they are as expected. We say $\mathbf{M} = (M, \pi)$ *divides* $\mathbf{M}' = (M', \pi')$ if there exists a sub $\circledast$-monoid $\mathbf{M}'' = (M'', \pi'')$ of $\mathbf{M}'$ and a surjective morphism from $\mathbf{M}''$ to $\mathbf{M}$.

A $\circledast$-monoid $\mathbf{M} = (M, \pi)$ is said to be finite if M is so. Note that, even for a finite $\circledast$-monoid, the product operation π has an infinitary description. It turns out that π can be captured using finitely presentable *derived operations*. Corresponding to a $\circledast$-monoid (M, π) there is an induced *$\circledast$-algebra* $\mathbf{M} = (M, \mathtt{id}, \cdot, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\kappa})$ where the operations are defined as following: for all $a, b \in M$, $a \cdot b = \pi(ab)$, $a^{\boldsymbol{\tau}} = \pi(a^{\omega})$, $a^{\boldsymbol{\tau}^*} = \pi(a^{\omega^*})$ and for all $\emptyset \neq E \subseteq M$, $E^{\boldsymbol{\kappa}} = \pi(E^{\eta})$. For a singleton set $\{m\}$, we write $m^{\boldsymbol{\kappa}} = \{m\}^{\boldsymbol{\kappa}}$. These derived operators satisfy certain natural axioms; see [4] for details. It has been established in [4] that an *arbitrary* finite $\circledast$-algebra $\mathbf{M} = (M, \mathtt{id}, \cdot, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\kappa})$ satisfying these natural axioms is induced by a unique $\circledast$-monoid $\mathbf{M} = (M, \pi)$. It is rather straightforward to define the notions of morphisms, subalgebras, direct-products as well as division for $\circledast$-algebras.

It follows from the definition of a $\circledast$-algebra $\mathbf{M} = (M, \mathtt{id}, \cdot, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\kappa})$ that $(M, \mathtt{id}, \cdot)$ is a *monoid*, that is the operation $\cdot$ is associative with identity $\mathtt{id}$. Note that, for all $m \in M$, $m \cdot \mathtt{id} = \mathtt{id} \cdot m = m$ and for all $\emptyset \neq E \subseteq M$, $E^{\boldsymbol{\kappa}} = (E \cup \{\mathtt{id}\})^{\boldsymbol{\kappa}}$. Further, $\mathtt{id}^{\boldsymbol{\tau}} = \mathtt{id}^{\boldsymbol{\tau}^*} = \mathtt{id}^{\boldsymbol{\kappa}} = \mathtt{id}$. As a result, in our definitions of $\circledast$-algebras later in the paper, we restrict the descriptions of derived operators to $M \setminus \{\mathtt{id}\}$. An *idempotent* is an element e where $e \cdot e = e$. For an element m, the *idempotent power* $m^!$ is the idempotent got by multiplying m a finite number of times. An idempotent power exists for all elements in a finite $\circledast$-monoid.

An *evaluation tree* over a word $u \in M^{\circledast} \backslash \{\varepsilon\}$ is a tree $\mathcal{T} = (T, h)$ such that every branch/path of $\mathcal{T}$ is of finite length and where every vertex in T is a factor of u, the root is u and $h : T \to M$ is a map such that:

- A leaf is a singleton letter $a \in M$ such that $h(a) = a$.
- Internal nodes have either two or ω or ω^* or $\mathbb{Q}$ many children.
- If w has children v_1 and v_2, then $w = v_1 v_2$ and $h(w) = h(v_1) \cdot h(v_2)$.
- If w has ω many children $\langle v_1, v_2, \dots \rangle$, then there is an idempotent e such that $e = h(v_i)$ for all $i \geq 1$, and $w = \prod_{i \in \omega} v_i$ and $h(w) = e^{\boldsymbol{\tau}}$.
- If w has ω^* many children $\langle \dots, v_{-2}, v_{-1} \rangle$, then there is an idempotent f such that $f = h(v_i)$ for all $i \leq -1$, and $w = \prod_{i \in \omega^*} v_i$ and $h(w) = f^{\boldsymbol{\tau}^*}$.
- If w has $\mathbb{Q}$ many children $\langle v_i \rangle_{i \in \mathbb{Q}}$, then $w = \prod_{i \in \mathbb{Q}} v_i$ where for the perfect shuffle f over an $E = \{a_1, \dots, a_k\} \subseteq M$, $h(v_i) = a_{f(i)}$, and $h(w) = E^{\boldsymbol{\kappa}}$.

The *value* of $\mathcal{T}$ is defined to be $h(u)$. It was shown in [4, Proposition 8 and 9] that every word u has an evaluation tree and the values of two evaluation trees of u are equal and they are equal to $\pi(u)$. Therefore, a $\circledast$-algebra defines the generalized associativity product $\pi : M^{\circledast} \to M$. The correspondence between finite $\circledast$-monoids and $\circledast$-algebras permits interchangeability; we exploit it implicitly.

A morphism from the *free* $\circledast$*-monoid* $\Sigma^{\circledast}$ to $\mathbf{M}$ is described (determined) by a map $h' : \Sigma \to M$; we simply write $h' : \Sigma \to \mathbf{M}$. With h' also denoting its pointwise extension $h' : \Sigma^{\circledast} \to M^{\circledast}$, given a word $u \in \Sigma^{\circledast}$, we can use the evalution tree over the word $h'(u) \in M^{\circledast}$ to obtain $\pi(h'(u)) \in M$. By further abuse of notation, $h' : \Sigma^{\circledast} \to M$ also denotes the morphism which sends u to $\pi(h'(u))$. We say that L is recognized by $\mathbf{M}$ if there exists a map/morphism $h' : \Sigma^{\circledast} \to \mathbf{M}^{\circledast}$ such that $L = h'^{-1}(h'(L))$. The fundamental result of [4] states that *regular languages* (MSO definable languages) are exactly those recognized by finite $\circledast$-monoids (equivalently $\circledast$-algebras). It is important to note that, every regular language L is associated a finite (canonical/minimal) syntactic $\circledast$-monoid which divides every $\circledast$-monoid that recognizes L. Further, it can be represented as a $\circledast$-algebra from a finite description of L.

Example 1. The $\circledast$-monoid $\mathrm{U}_1 = (\{\mathtt{id}, 0\}, \pi)$ and its induced $\circledast$-algebra are shown on the left and right respectively.

$$\pi(u) = \begin{cases} \mathtt{id} & \text{if } u \in \{\mathtt{id}\}^{\circledast} \\ 0 & \text{otherwise} \end{cases}$$

	id	0	$\boldsymbol{\tau}$	$\boldsymbol{\tau}^*$
id	id	0	id	id
0	0	0	0	0

$$S^{\boldsymbol{\kappa}} = \begin{cases} \mathtt{id} & \text{if } S = \{\mathtt{id}\} \\ 0 & \text{otherwise} \end{cases}$$

Let $\Sigma = \{a, b\}$ and L be the set of words which contain an occurence of letter a. It is easy to see that the map $h : \Sigma \to \mathrm{U}_1$ sending $h(a) = 0, h(b) = \mathtt{id}$ recognizes L as $L = h^{-1}(0)$. In fact, U_1 is the syntactic $\circledast$-monoid of L.

Example 2. Consider the $\circledast$-algebra Gap$= (\{\mathtt{id}, [\,], (\,], [\,), (\,), g\}, \mathtt{id}, \cdot, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\kappa})$. We let $\Sigma = \{a\}$ and define the map $h : \Sigma \to \text{Gap}$ as $h(a) = [\,]$. The resulting morphism maps a word u to $h(u) = g$ iff the word u admits a gap; that is a cut with no maximum and its complement has no minimum. Other words are mapped to their right 'ends-type': for instance, $h(u) = [\ )$ iff $dom(u)$ has a minimum and no maximum. For a word $v = a^{\omega} a^{\omega^*}$, the pointwise extension $v' = h(v) = [\ ]^{\omega}[\ ]^{\omega^*}$. An example evaluation tree $\mathcal{T}$ for v' consists of root with two children. The left (resp. right) child has ω (resp. ω^*) many children $[\ ]$ and has value $[\ ]^{\boldsymbol{\tau}}$ (resp. $[\ ]^{\boldsymbol{\tau}^*}$). As a result, the value of $\mathcal{T}$ is $[\ ]^{\boldsymbol{\tau}} \cdot [\ ]^{\boldsymbol{\tau}^*} = [\ ) \cdot (\ ] = g$.

$\cdot$	$[\ ]$	$[\ )$	$(\ ]$	$(\ )$	g	$\boldsymbol{\tau}$	$\boldsymbol{\tau}^*$
$[\ ]$	$[\ ]$	$[\ )$	$[\ ]$	$[\ )$	g	$[\ )$	$(\ ]$
$[\ )$	$[\ ]$	$[\ )$	g	g	g	$[\ )$	$(\ )$
$(\ ]$	$(\ ]$	$(\ )$	$(\ ]$	$(\ )$	g	$(\ )$	$(\ ]$
$(\ )$	$(\ ]$	$(\ )$	g	g	g	g	g
g	g	g	g	g	g	g	g

$$S^{\boldsymbol{\kappa}} = \begin{cases} \texttt{id} & \text{if } S = \{\texttt{id}\} \\ g & \text{otherwise} \end{cases}$$

We can characterize $\circledast$-monoids using equational *identities*. For example, $\mathbf{M}$ is a *commutative* $\circledast$-monoid if and only if $\mathbf{M}$ satisfies the equation $x \cdot y = y \cdot x$. This means that the equation holds for any assignment of elements in the $\mathbf{M}$ to the variables x and y. We say $\mathbf{M}$ is aperiodic if it satisfies the profinite identity $x = x \cdot x^!$. Like in the case of monoids, the set of $\circledast$-monoids satisfying a set of equations are closed under subsemigroup, division and *direct product* [5].

The *block product* of $\circledast$-monoids $\mathbf{M}$ and $\mathbf{N}$, is denoted by $\mathbf{M} \Box \mathbf{N}$ and is the semidirect product of $\mathbf{M}$ and $\mathbf{K} = \mathbf{N}^{M \times M}$ with respect to the *canonical* left and right 'action' of $\mathbf{M}$ on $\mathbf{K}$. The details are given in [1]. The *block product principle* characterizes languages defined by block product of $\circledast$-monoids. Towards this, fix a map $h : \Sigma \to \mathbf{M} \Box \mathbf{N}$ such that $h(a) = (m_a, f_a)$ where $m_a \in M$ and $f_a : M \times M \to N$. The map $h_1 : \Sigma \to \mathbf{M}$ setting $h_1(a) = m_a$ defines a morphism $h_1 : \Sigma^{\circledast} \to M$. We define the *transducer* $\sigma : \Sigma^{\circledast} \to (M \times \Sigma \times M)^{\circledast}$ as follows: let $u \in \Sigma^{\circledast}$ with domain α. The word $u' = \sigma(u)$ has domain α and for a position $x \in \alpha$, $u'(x) = (h_1(u_{<x}), u(x), h_1(u_{>x}))$. Here $u_{<x}$ (resp. $u_{>x}$)) is the subword of u on positions strictly less (resp. greater) than x.

Proposition 1 (Block Product Principle [1]). *Let $L \subseteq \Sigma^{\circledast}$ be recognized by $h : \Sigma \to \mathbf{M} \Box \mathbf{N}$ Then L is a boolean combination of languages of the form L_1 and $\sigma^{-1}(L_2)$ where L_1 and L_2 are recognized by $\mathbf{M}$ and $\mathbf{N}$ respectively and $\sigma : \Sigma^{\circledast} \to (M \times \Sigma \times M)^{\circledast}$ is a state-based transducer.*

3 Small fragments of FO

In this section, we focus on two particularly small fragments of first-order logic interpreted over countable words. First-order logic uses variables $x, y, z, \ldots$ which are interpreted as positions in the domain of a word. The syntax of *first-order logic* (FO) is: $x < y \mid a(x) \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2 \mid \neg \phi \mid \exists x\ \phi$, for all $a \in \Sigma$.

We skip the natural semantics. A language L of countable words is said to be FO-definable if there exists an FO-sentence ϕ such $L = \{u \in \Sigma^{\circledast} \mid u \models \phi\}$.

Recall that the classical Schutzenberger-McNaughton-Papert theorem characterizes FO-definabilty of a regular language of finite words in terms of aperiodicity of its finite syntactic monoid. The survey [6] presents similar decidable characterizations of several interesting small fragments of FO-logic such as FO^1, FO^2, $B(\exists^*)$ – boolean closure of the existential first-order logic. It is known [7] that, over finite *as well as countable words*, FO $= \mathrm{FO}^3$. As mentioned in the introduction, over countable words, we already have decidable algebraic characterizations of FO^3 from [5] and FO^2 from [9]. Here we identify decidable algebraic characterizations, over countable words, for FO^1 and $B(\exists^*)$.

3.1 FO with single variable

The fragment FO^1 has access to only one variable. We recall that over finite words a regular language is FO^1-definable iff its syntactic monoid is commutative and idempotent. We henceforth focus our attention to FO^1 on countable words.

Clearly, FO^1 can recognize all words with a particular letter. With a single variable the logic cannot talk about order of letters or count the number of occurrence of a letter. This gives an intuition that the syntactic $\circledast$-monoid of a language definable in FO^1 is commutative and idempotent.

We say that a $\circledast$-algebra $\mathbf{M} = (M, \mathtt{id}, \cdot, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\kappa})$ is *shuffle-trivial* if it satisfies the equational identity: $\{x_1, \ldots, x_p\}^{\boldsymbol{\kappa}} = x_1 \cdot x_2 \cdot \ldots \cdot x_p$.

Then $\mathbf{M}$ is commutative: $x \cdot y = \{x, y\}^{\boldsymbol{\kappa}} = \{y, x\}^{\boldsymbol{\kappa}} = y \cdot x$. Moreover, every element of $\mathbf{M}$ is a *shuffle-idempotent*: for all $m \in M, m^{\boldsymbol{\kappa}} = m$. It is a consequence of the axioms of a $\circledast$-algebra that a shuffle-idempotent is an idempotent.

Theorem 1. *Let $L \subseteq \Sigma^{\circledast}$ be a regular language. The following are equivalent.*

1. *L is recognized by some finite shuffle-trivial $\circledast$-algebra.*
2. *L is a boolean combination of languages of the form $B^{\circledast}$ where $B \subseteq \Sigma$.*
3. *L is definable in FO^1.*
4. *L is recognized by direct product of U_1s.*
5. *The syntactic $\circledast$-algebra of L is shuffle-trivial.*

3.2 Boolean closure of existential FO

Let us first recall the characterization of $B(\exists^*)$ - the boolean closure of existential FO over finite words. This is precisely the content of the theorem due to Simon [14]. The usual presentation of Simon's theorem refers to *piecewise testable languages* which are easily seen to be equivalent to $B(\exists^*)$-definable languages. *Simon's theorem* states that a regular language of finite words is $B(\exists^*)$-definable iff its syntactic monoid is J-trivial. We refer to [11] for a detailed study of Green's relations and its use in the proof of Simon's theorem.

The original proof of Simon's theorem uses the congruence $\sim_n$, parametrized by $n \in \mathbb{N}$, on finite words Σ^*: for $u, v \in \Sigma^*$, $u \sim_n v$ if u and v have the same set of subwords of length less than or equal to n. Note that $\sim_n$ has finite index.

We fix $n \in \mathbb{N}$ and work with $\sim_n$ defined on countable words $\Sigma^{\circledast}$: for $u, v \in \Sigma^{\circledast}$, $u \sim_n v$ if u and v have the same set of subwords of length less than or equal to n. It is immediate that $\sim_n$ is an equivalence relation on $\Sigma^{\circledast}$ of finite index. We let $S_n = \Sigma^{\circledast} / \sim_n$ denote the finite set of $\sim_n$-equivalence classes. For a word w, $[w]_n$ denotes the $\sim_n$-equivalence class which contains w.

Lemma 1. *There is a natural well-defined product operation $\pi : S_n^{\circledast} \to S_n$ as follows: $\pi\Big(\prod_{i\in\alpha}[w_i]_n\Big) = \big[\prod_{i\in\alpha} w_i\big]_n$. This operation π satisfies the generalized associativity property. As a result, $\mathbf{S_n} = (S_n, \mathtt{id} = [\varepsilon]_n, \pi)$ is a $\circledast$-monoid.*

Note that the lemma implies that $h_n : \Sigma^{\circledast} \to \mathbf{S_n}$ mapping w to $[w]_n$ is a morphism of $\circledast$-*monoids*.

We say that a $\circledast$-algebra is *shuffle-power-trivial* if it satisfies the (profinite) identity: $\{x_1, \ldots, x_p\}^{\boldsymbol{\kappa}} = (x_1 \cdot x_2 \cdot \ldots \cdot x_p)^!$. Note that, every idempotent of such a $\circledast$-algebra is a shuffle-idempotent: $x^! = x$ implies $x^{\boldsymbol{\kappa}} = x$.

Theorem 2. *Let $L \subseteq \Sigma^{\circledast}$ be a regular language. The following are equivalent.*

1. *L is recognized by a finite shuffle-power-trivial $\circledast$-algebra.*
2. *L is recognized by the quotient morphism $h_n : \Sigma^{\circledast} \to \mathbf{S_n}$ for some n.*
3. *L is definable in $B(\exists^*)$.*
4. *The syntactic $\circledast$-algebra of L is shuffle-power-trivial.*

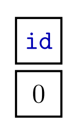
4 First Order Logic with infinitary quantifiers

Our results in the previous section resemble very closely the corresponding results over finite words. This can be attributed to the limited capability of the operators $\boldsymbol{\tau}$, $\boldsymbol{\tau}^*$ and $\boldsymbol{\kappa}$ in the $\circledast$-monoids we witnessed. As mentioned in the Introduction, FO cannot define the language of infinite number of a's. An existential quantifier is a threshold counting quantifier - it says there exists at least one position satisfying a property. Using multiple such first-order quantifiers, FO can count up to any finite constant but not more. Over countable words, it is natural to ask for stronger threshold quantifiers. We introduce natural infinite extensions of the existential quantifier. These quantifiers can distinguish ordinals in the infinite.

We define $\mathcal{I}_0$ to be the set of all non-empty finite orderings. For any number $n \in \mathbb{N}$, we define the set $\mathcal{I}_n$ to be the set of all orderings of the form $\sum_{i\in\mathbb{Z}} \alpha_i$ where $\alpha_i \in \mathcal{I}_{n-1} \cup \{\varepsilon\}$ and is closed under finite sum. We define the Infinitary rank (or simply *rank*) of a linear ordering α (denoted by $\mathcal{I}$-*rank*(α)) as the least n (if it exists) where $\alpha \in \mathcal{I}_n$. If there is no such n we say that the rank is infinite. For example, $\mathcal{I}\text{-}rank(\omega) = \mathcal{I}\text{-}rank(\omega+\omega) = \mathcal{I}\text{-}rank(\omega^*+\omega) = 1$, $\mathcal{I}\text{-}rank(\omega^2) = \mathcal{I}\text{-}rank(\omega^2+\omega^*) = 2$, and the rank of rational numbers is infinite.

We introduce the logic FO$[\infty]$ extending FO with *infinitary quantifiers* : $\exists^{\infty_0} x\ \varphi \mid \exists^{\infty_1} x\ \varphi \mid \ldots \mid \exists^{\infty_n} x\ \varphi \mid \ldots$ for all $n \in \mathbb{N}$.

Note that all the variables are first order. The semantics of the infinitary quantifier $\exists^{\infty_n} x$ for an $n \geq 0$ is: for a word w and an assignment s, we say $w, s \models \exists^{\infty_n} x\ \varphi$ if there exists a subordering $X \subseteq dom(w)$ such that $\mathcal{I}\text{-}rank(X) = n$ and $w, s[x = i] \models \varphi$ for all $i \in X$. For example, $\exists^{\infty_0} x\ \varphi$ is equivalent to $\exists x\ \varphi$ since both formulas are true if and only if there is at least one satisfying assignment $x = s$.



Δ_0-chain

The logic $\mathrm{FO}[(\infty_j)_{j\leq n}]$ denote the fragment containing only the infinitary quantifiers $\exists^{\infty_j} x$ for all $j \leq n$. Clearly the following relationship is maintained among the logics:

$$\mathrm{FO} = \mathrm{FO}[(\infty_j)_{j\leq 0}] \subseteq \mathrm{FO}[(\infty_j)_{j\leq 1}] \subseteq \mathrm{FO}[(\infty_j)_{j\leq 2}] \subseteq \ldots$$

We also denote by $\mathrm{FO}^1[(\infty_j)_{j\leq n}]$ the corresponding one variable fragment of $\mathrm{FO}[(\infty_j)_{j\leq n}]$.

Example 3. The formula $\exists^{\infty_1} x\ a(x)$ denotes the set of all countable words with infinitely many positions labelled a. Since FO cannot express this, it shows $\mathrm{FO} \subsetneq \mathrm{FO}[(\infty_j)_{j\leq 1}]$.

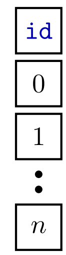


Fig. 1: Δ_n-chain

For an $n \geq 0$, we define the $\circledast$-algebra Δ_n-chains as: $(\{\mathtt{id}, 0, 1, \ldots, n\}, \mathtt{id}, \boldsymbol{\cdot}, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\kappa})$ where for all $0 \leq i \leq j \leq n$, $i \boldsymbol{\cdot} j = j \boldsymbol{\cdot} i = \max(i, j) = j$ and for all $0 \leq k < n$, $k^{\boldsymbol{\tau}} = k^{\boldsymbol{\tau}^*} = k + 1$ and $n^{\boldsymbol{\tau}} = n^{\boldsymbol{\tau}^*} = n$. That is, $k^{\boldsymbol{\tau}} = k^{\boldsymbol{\tau}^*} = \min(k+1, n)$ Moreover, $\mathtt{id}^{\boldsymbol{\kappa}} = \mathtt{id}$ and $S^{\boldsymbol{\kappa}} = n$ for any S where $S \setminus \{\mathtt{id}\} \neq \emptyset$.

$$\Delta_n ::= (\{\mathtt{id}, 0, 1, \ldots, n\}, \{i, j\} \overset{\boldsymbol{\cdot}}{\longmapsto} \max(i, j), i \overset{\boldsymbol{\tau}}{\longmapsto} \min(i{+}1, n), i \overset{\boldsymbol{\tau}^*}{\longmapsto} \min(i{+}1, n), S \overset{\boldsymbol{\kappa}}{\longmapsto} n)^2$$

Note that the syntactic $\circledast$-algebra for the language defined by $\exists^{\infty_n} x\ a(x)$ is Δ_n.

4.1 FO[∞] with single variable

In this section we show that languages recognized by Δ_n are definable in $\mathrm{FO}^1[(\infty_j)_{j\leq n}]$. It is easy to observe that the direct product of Δ_n recognize exactly those languages definable in the one variable fragment.

Theorem 3. *Languages recognized by direct product of Δ_n are exactly those definable in $\mathrm{FO}^1[(\infty_j)_{j\leq n}]$.*

Proof. We first show that languages recognized by Δ_n are definable in $\mathrm{FO}^1[(\infty_j)_{j\leq n}]$. Let $h\colon \Sigma^{\circledast} \to \Delta_n$ be a morphism. It suffices to show that for any element $m \in \Delta_n$, $h^{-1}(m)$ is definable in $\mathrm{FO}^1[(\infty_j)_{j\leq n}]$. In the rest of the discussion we adopt the convention that $\mathtt{id} < 0$. Let $\uparrow m$ denote the set $\{m' \mid m' \geq m\}$. Note that for an $m < n$, $h^{-1}(m) = h^{-1}(\uparrow m) \setminus h^{-1}(\uparrow(m+1))$ and $h^{-1}(n) = h^{-1}(\uparrow n)$. Therefore, it is sufficient to show that $h^{-1}(\uparrow m)$ is definable in $\mathrm{FO}^1[(\infty_j)_{j\leq n}]$. For each $m \in \Delta_n$, we define the language $L(m)$ as
$\{w \mid$ there exists a letter a in w such that $h(a) = j \neq \mathtt{id}$ and either $j \geq m$ or there is a set of positions α labelled a such that $\mathcal{I}\text{-}\mathit{rank}(\alpha) = j'$ and $j + j' \geq m\}$
The following $\mathrm{FO}^1[(\infty_j)_{j\leq n}]$ sentence defines the language $L(m)$.

$$\bigvee_{a\in\Sigma,\ h(a)\geq m} \exists x\ a(x) \quad \vee \quad \bigvee_{a\in\Sigma,\ 0\leq h(a)<m} \exists^{\infty_{m-h(a)}} x\ a(x)$$

[2] As mentioned in the Preliminaries, we restrict the descriptions of derived operators to $\Delta_n \setminus \{\mathtt{id}\}$

We show that $L(m) = h^{-1}(\uparrow m)$ by induction on m. The base case holds since $\uparrow\texttt{id} = \Delta_n$, $h^{-1}(\uparrow\texttt{id}) = \Sigma^{\circledast}$ and $L(\texttt{id}) = \Sigma^{\circledast}$. To prove the induction hypothesis assume the claim holds for all $j < m$. Consider a word w. By a second induction on the height of an evaluation tree (T, h) for w we show for all words $v \in T$, $v \in h^{-1}(\uparrow m)$ if and only if $v \in L(m)$. In each of the following cases we assume that the children of the node (if they exist) satisfy the second induction hypothesis.

1. Case v is a letter: The hypothesis clearly holds
2. Case v is a concatenation of two words v_1 and v_2: There are two cases to consider - $\{v_1, v_2\} \cap h^{-1}(\uparrow m) \neq \emptyset$ or not. In the first case, let for an $i \in \{1, 2\}$ we have $h(v_i) \geq m$ and $v_i \in L(m)$. Clearly $h(v) = h(v_1 v_2) \geq m$ and $v \in L(m)$. For the second case, let us assume $h(v_1) = i$ and $h(v_2) = j$ such that $i \leq j < m$ and both $v_1, v_2 \notin L(m)$. From the definition of Δ_n, it follows that $h(v) = h(v_1 v_2) = j$. Let the a-labelled suborderings in v_1 and v_2 be α_1 and α_2 respectively where $\mathcal{I}\text{-}\mathit{rank}(\alpha_1) \leq \mathcal{I}\text{-}\mathit{rank}(\alpha_2) = j'$. It follows from the definition that $\mathcal{I}\text{-}\mathit{rank}(\alpha_1 + \alpha_2) = j'$ and therefore $v \notin L(m)$.
3. Case v is an omega sequence of words $\langle v_1, v_2, \ldots, \rangle$ such that $h(v_i) = k$, for all i, and k is an idempotent (in Δ_n all elements are idempotents): Firstly, if $k \geq m$ and $v_i \in L(m)$ then clearly $h(v) \geq m$ and $v \in L(m)$. The non-trivial case is $k = m - 1$. From the second induction hypothesis $v_i \notin L(m)$ for all i. From the definition of Δ_n, $h(v) = k^{\boldsymbol{\tau}} = m$. We need to show that $v \in L(m)$. By first induction hypothesis, each v_i has a letter a_i and an a_i-labelled set of positions α_i such that $h(a_i) + \mathcal{I}\text{-}\mathit{rank}(\alpha_i) = k$. Since $|\Sigma|$ is finite, there is a letter a occurring in omega many factors. Hence the a-labelled set of positions α in v satisfies $h(a) + \mathcal{I}\text{-}\mathit{rank}(\alpha) = k+1$ or in other words $v \in L(m)$.
4. Case v is an omega* sequence: This case is symmetric to the above case.
5. Case v is a perfect shuffle, $h(v) = S^{\boldsymbol{\kappa}}$: It is easy to see that the induction hypothesis holds if $S = \{\texttt{id}\}$. So, assume $S \cap \{\texttt{id}\} \neq \emptyset$. Hence $h(v) = n$. Since, there are rational number of children u where $h(u) \neq \texttt{id}$, there is a letter a such that a-labelled set of positions in v has infinite rank or $v \in L(n)$.

The other direction of the proof follows from the fact that a one variable quantifier free formula is essentially a disjunction of letter predicates and therefore the boolean combination of formulas can be recognized by direct products of Δ_k.

4.2 The general $\mathbf{FO}[\infty]$ logic

In this section, we consider the full logic $\mathrm{FO}[(\infty_j)_{j \leq n}]$ and observe that they define exactly those languages recognized by block products of Δ_n.

Theorem 4. *The languages defined by* $\mathrm{FO}[(\infty_j)_{j \leq n}]$ *are exactly those recognized by finite block products of* Δ_n*. Moreover, the languages defined by* $\mathrm{FO}[\infty]$ *are exactly those recognized by finite block products of* $\{\Delta_n \mid n \in \mathbb{N}\}$.

Proof. We first show that languages recognizable by finite block products of Δ_n are definable in $\mathrm{FO}[(\infty_j)_{j \leq n}]$. The proof is via induction on the number of Δ_n in an iterated block product. The base case follows from Theorem 3.

For the inductive step, consider a morphism $h\colon \Sigma^{\circledast} \to M\Box\Delta_n$. Let $h_1 : \Sigma^{\circledast} \to M$ be the induced morphism to M, and let σ be the associated transducer. By the block product principle (see Proposition 1), any language recognized by h is a boolean combination of languages $L_1 \subseteq \Sigma^{\circledast}$ recognized by M and $\sigma^{-1}(L_2)$ where $L_2 \subseteq (M \times \Sigma \times M)^{\circledast}$ is recognized by Δ_n. By induction hypothesis, L_1 is $\mathrm{FO}[(\infty_j)_{j\leq n}]$ definable. By the base case L_2 is $\mathrm{FO}[(\infty_j)_{j\leq n}]$ definable but over the alphabet $M \times \Sigma \times M$. To complete the proof, one needs to show for any word $w \in \Sigma^{\circledast}$ and assignment s, and for any $\mathrm{FO}[(\infty_j)_{j\leq n}]$ formula φ over the alphabet $M \times \Sigma \times M$, there exists a $\mathrm{FO}[(\infty_j)_{j\leq n}]$ formula $\hat{\varphi}$ over the alphabet Σ such that $w, s \models \hat{\varphi}$ if and only if $\sigma(w), s \models \varphi$. For instance, suppose $\varphi = \exists^{\infty_i} x\ (m_1, c, m_2)(x)$, and inductively ϕ_{m_1} (resp. ϕ_{m_2}) are $\mathrm{FO}[(\infty_j)_{j\leq n}]$ formula characterizing words over $\Sigma^{\circledast}$ that are mapped by h_1 to m_1 (resp. m_2). Then $\hat{\varphi}$ is $\exists^{\infty_i} x\ (\phi_{m_1}|_{<x} \wedge c(x) \wedge \phi_{m_2}|_{>x})$, where $\phi_{m_1}|_{<x}$ is the formula ϕ_{m_1} with all its variables relativised to less than the variable x. This way, one proves that $\sigma^{-1}(L_2)$ is $\mathrm{FO}[(\infty_j)_{j\leq n}]$ definable. This completes the proof of this direction.

The other direction of the proof is a standard generalization of the proof of equivalence of FO and the block product closure of Δ_0 given in [1, Theorem 2]. The block product principle allows us to “simulate” infinitary quantifiers using block products of Δ_n and vice-versa. We can then inductively recognize languages defined by formulas using iterated block products.

We claim that both first order logic with cuts (FO[cut]) and weak monadic second order logic (WMSO) can define the languages definable in $\mathrm{FO}[\infty]$.

Theorem 5. $\mathrm{FO}[\infty] \subseteq \mathrm{FO}[\mathrm{cut}] \cap \mathrm{WMSO}$ [3]

5 No Finite Basis Theorems

The main goal of this section is to prove that $\mathrm{FO}[\infty]$, FO[cut] and $\mathrm{FO}[\mathrm{cut}] \cap \mathrm{WMSO}$ over countable words do not admit a block product based characterization which uses only a *finite* set of $\circledast$-monoids. This is in stark contrast with the result in [1] which shows that a language of countable words is FO-definable iff it is recognized by a strong iteration of block product of copies of Δ_0 (alternately called U_1). This is abbreviated by saying that FO has a block-product based characterization using a basis which contains the single $\circledast$-monoid Δ_0. Notice that, it follows from the results in the previous section that $\mathrm{FO}[\infty]$ admits a block product based characterization using the *natural infinite basis* $\{\Delta_n\}_{n\in\mathbb{N}}$.

Fix a finite $\circledast$-algebra $\mathbf{M} = (M, \mathtt{id}, \cdot, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\kappa})$. For every $n \in \mathbb{N}$, we define the operation $\gamma_n : M \to M$ which maps x to x^{γ_n}. The inductive definition of γ_n is as follows (recall idempotent power): $x^{\gamma_0} = x^!$ and $x^{\gamma_n} = ((x^{\gamma_{n-1}})^{\boldsymbol{\tau}}(x^{\gamma_{n-1}})^{\boldsymbol{\tau}^*})^!$.

Lemma 2. *For each $m \in M$, there exists n such that $\forall n' \geq n, m^{\gamma_n} = m^{\gamma_{n'}}$.*

We now define the *gap-nesting-length* of $\mathbf{M}$ (in notation, $\mathrm{gnlen}(\mathbf{M})$) to be the smallest n such that for all $m \in M$, $m^{\gamma_n} = m^{\gamma_{n+1}}$. It follows from the previous

[3] Here, $\mathrm{FO}[\infty]$, FO[cut], WMSO denote the languages defined by the respective logic.

lemma that a finite $\circledast$-algebra has a finite gap-nesting-length. It is a simple computation that, for each k, $\mathrm{gnlen}(\Delta_k) = k$. The following main technical lemma is the key to our no-finite-basis theorems.

Lemma 3. *For finite aperiodic*[4] *$\circledast$-algebras* $\mathbf{M}$ *and* $\mathbf{N}$ *,*

1. *We have,* $\mathrm{gnlen}(\mathbf{M}\Box\mathbf{N}) \leq \max(\mathrm{gnlen}(\mathbf{M}), \mathrm{gnlen}(\mathbf{N}))$.
2. *If* $\mathbf{M}$ *divides* $\mathbf{N}$ *then* $\mathrm{gnlen}(\mathbf{M}) \leq \mathrm{gnlen}(\mathbf{N})$.

Corollary 1. $\mathrm{FO}[(\infty_j)_{j\leq n}] \subsetneq \mathrm{FO}[(\infty_j)_{j\leq n+1}]$.

Proof. By Theorem 4, the syntactic $\circledast$-algebra $\mathbf{M}$ of any $\mathrm{FO}[(\infty_j)_{j\leq n}]$-definable language divides a block product of copies of Δ_n. By Lemma 3 and the fact that $\mathrm{gnlen}(\Delta_k) = n$, $\mathrm{gnlen}(\mathbf{M}) \leq n$. Note that, Δ_{n+1} is the syntactic $\circledast$-algebra for the language L defined by the $\mathrm{FO}[(\infty_j)_{j\leq n+1}]$ formula $\exists^{\infty_{n+1}} x\ a(x)$. As $\mathrm{gnlen}(\Delta_{n+1}) = n+1$, it follows that L cannot be defined in $\mathrm{FO}[(\infty_j)_{j\leq n}]$.

Theorem 6. *There is no finite basis for a block product based characterization for any of these logical systems* $\mathrm{FO}[\infty], \mathrm{FO}[\mathrm{cut}], \mathrm{FO}[\mathrm{cut}] \cap \mathrm{WMSO}$.

Proof. Fix one of the logics $\mathcal{L}$ mentioned in the statement of the theorem. It follows from Theorem 5 and the algberaic chacterization [5] of $\mathrm{FO}[\mathrm{cut}]$ that the syntactic $\circledast$-algebras of $\mathcal{L}$-definable languages are aperiodic. Now suppose, for contradiction, let $\mathcal{L}$ admit a finite basis B of aperiodic $\circledast$-algebras for its block product based chacterization. Since B is finite, there exists $n \in \mathbb{N}$ such that for all $\circledast$-algebras $\mathbf{M}$ in B, $\mathrm{gnlen}(\mathbf{M}) \leq n$. It follows by Lemma 3 that the syntactic $\circledast$-algebra $\mathbf{N}$ of *every* $\mathcal{L}$-definable language has the property $\mathrm{gnlen}(\mathbf{N}) \leq n$.

Now consider the language L defined by the $\mathrm{FO}[\infty]$ sentence $\phi = \exists^{\infty_{n+1}} x\ a(x)$. By Theorem 5, L is $\mathcal{L}$-definable. Hence, the gap-nesting-length of the syntactic $\circledast$-algebra K of L is less than or equal to n. However, K is simply Δ_{n+1} and $\mathrm{gnlen}(\Delta_{m+1}) = n+1$. This leads to a contradiction.

References

1. Adsul, B., Sarkar, S., Sreejith, A.V.: Block products for algebras over countable words and applications to logic. In: 34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019, Vancouver, BC, Canada, June 24-27, 2019. pp. 1–13. IEEE (2019). https://doi.org/10.1109/LICS.2019.8785669, `https://doi.org/10.1109/LICS.2019.8785669`
2. Baudisch, A., Seese, D., Tuschik, H.P., Weese, M.: Decidability and Generalized Quantifiers. Akademie Verlag, Berlin (1980)
3. Bès, A., Carton, O.: Algebraic characterization of FO for scattered linear orderings. In: Bezem, M. (ed.) Computer Science Logic, 25th International Workshop / 20th Annual Conference of the EACSL, CSL 2011, September 12-15, 2011, Bergen, Norway, Proceedings. LIPIcs, vol. 12, pp. 67–81. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2011)

[4] This means the underlying monoid of a $\circledast$-algebra is aperiodic

4. Carton, O., Colcombet, T., Puppis, G.: An algebraic approach to MSO-definability on countable linear orderings. J. Symb. Log. **83**(3), 1147–1189 (2018). https://doi.org/10.1017/jsl.2018.7, https://doi.org/10.1017/jsl.2018.7
5. Colcombet, T., Sreejith, A.V.: Limited set quantifiers over countable linear orderings. In: Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Proceedings, Part II. pp. 146–158 (2015)
6. Diekert, V., Gastin, P., Kufleitner, M.: A survey on small fragments of first-order logic over finite words. Int. J. Found. Comput. Sci. **19**(3), 513–548 (2008). https://doi.org/10.1142/S0129054108005802, https://doi.org/10.1142/S0129054108005802
7. Gabbay, D.M., Hodkinson, I., Reynolds, M.: Temporal Logic: Mathematical Foundations and Computational Aspects, Volume 1. Oxford University Press, Oxford (1994)
8. Gradel, E., Otto, M., Rosen, E.: Two-variable logic with counting is decidable. In: Proceedings of Twelfth Annual IEEE Symposium on Logic in Computer Science. pp. 306–317 (1997). https://doi.org/10.1109/LICS.1997.614957
9. Manuel, A., Sreejith, A.V.: Two-variable logic over countable linear orderings. In: 41st International Symposium on Mathematical Foundations of Computer Science, MFCS 2016. pp. 66:1–66:13 (2016). https://doi.org/10.4230/LIPIcs.MFCS.2016.66, https://doi.org/10.4230/LIPIcs.MFCS.2016.66
10. Pin, J.E.: Handbook of formal languages, vol. 1. chap. Syntactic Semigroups, pp. 679–746. Springer-Verlag (1997), http://dl.acm.org/citation.cfm?id=267846.267856
11. Pin, J.É.: Mathematical foundations of automata theory (2020)
12. Rosenstein, J.G.: Linear orderings. Academic Press New York (1981)
13. Shelah, S.: The monadic theory of order. vol. 102, pp. 379–419 (1975)
14. Simon, I.: Piecewise testable events. In: Brakhage, H. (ed.) Automata Theory and Formal Languages. pp. 214–222. Springer Berlin Heidelberg (1975)
15. Straubing, H.: Finite automata, formal logic, and circuit complexity. Birkhauser Verlag, Basel, Switzerland (1994)
16. Straubing, H., Thérien, D., Thomas, W.: Regular languages defined with generalized quantifiers. In: Lepistö, T., Salomaa, A. (eds.) Automata, Languages and Programming. pp. 561–575. Springer Berlin Heidelberg, Berlin, Heidelberg (1988)
17. Straubing, H., Weil, P.: Varieties. CoRR **abs/1502.03951** (2015), http://arxiv.org/abs/1502.03951
18. Thomas, W.: Handbook of formal languages, vol. 3. chap. Languages, Automata, and Logic, pp. 389–455. Springer-Verlag New York, Inc. (1997), http://dl.acm.org/citation.cfm?id=267871.267878