Theory Comput Syst
DOI 10.1007/s00224-017-9803-8

Syntactic Complexity of Regular Ideals

Janusz A. Brzozowski[1] · **Marek Szykuła**[2] · **Yuli Ye**[3,4]

Abstract The state complexity of a regular language is the number of states in a minimal deterministic finite automaton accepting the language. The syntactic complexity of a regular language is the cardinality of its syntactic semigroup. The syntactic complexity of a subclass of regular languages is the worst-case syntactic complexity taken as a function of the state complexity n of languages in that class. We prove that n^{n-1}, $n^{n-1}+n-1$, and $n^{n-2}+(n-2)2^{n-2}+1$ are tight upper bounds on the syntactic complexities of right ideals and prefix-closed languages, left ideals and suffix-closed languages, and two-sided ideals and factor-closed languages, respectively. Moreover, we show that the transition semigroups meeting the upper bounds for all three types of ideals are unique, and the numbers of generators (4, 5, and 6, respectively) cannot be reduced.

✉ Marek Szykuła
msz@cs.uni.wroc.pl

Janusz A. Brzozowski
brzozo@uwaterloo.ca

Yuli Ye
yuli.ye@gmail.com

[1] David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

[2] Institute of Computer Science, University of Wrocław, Joliot-Curie 15, 50-383 Wrocław, Poland

[3] Department of Computer Science, University of Toronto, Toronto, ON, M5S 3G4, Canada

[4] Present address: Wish.com, San Francisco, CA 94111, USA

Published online: 04 August 2017

Keywords Factor-closed · Left ideal · Prefix-closed · Regular language · Right ideal · Suffix-closed · Syntactic complexity · Transition semigroup · Two-sided ideal · Upper bound

1 Introduction

Formal definitions of the concepts introduced in this section are given in Section 2. For now we assume that the reader is familiar with basic properties of regular languages and finite automata as covered in [27, 32], for example.

There are two fundamental congruence relations in the theory of regular languages: the Nerode (right) congruence [25], and the Myhill congruence [24]. In both cases, a language is regular if and only if it is a union of congruence classes of a congruence of finite index. The Nerode congruence leads to the definitions of left quotients of a language and the minimal deterministic finite automaton (DFA) recognizing the language, and the Myhill congruence, to the definitions of the syntactic semigroup of the language.

The *state complexity* of a language is the number of states in a minimal DFA recognizing the language. This concept has been studied extensively; for surveys and references see [2, 33]. The *syntactic complexity* of a regular language is the cardinality of its *syntactic semigroup*, which is isomorphic to the *transition semigroup* of a minimal DFA recognizing the language [29], where the transition semigroup is the semigroup of transformations of the set of states of the DFA induced by non-empty words.

Syntactic complexity does not refine state complexity, for there exist languages with the same syntactic complexity but different state complexities. However, it often helps to distinguish among languages with the same state complexity. For example, the DFAs in Fig. 1 all have the same alphabet, are all minimal, and all have state complexity 3. However, the syntactic complexity of $\mathcal{D}_1$ is 3, that of $\mathcal{D}_2$ is 9, and that of $\mathcal{D}_3$ is 27.

The problem we study in this paper is the following: Given a language belonging to a subclass of the class of regular languages – for example, the subclass of finite languages or prefix-free languages (prefix-codes) – what is the maximal size of the

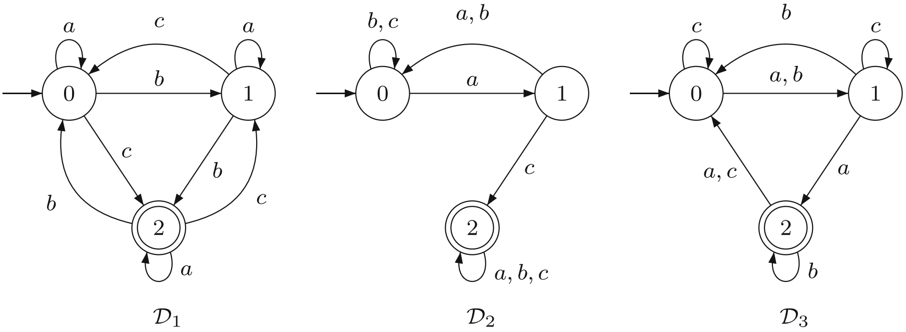

Fig. 1 DFAs with various syntactic complexities

syntactic semigroup of that language? Equivalently, given a minimal DFA of a language in the subclass, what is the maximal size of the transition semigroup of the DFA? A secondary problem is to find the minimal size of a set of generators for the maximal semigroup.

Syntactic complexity has been studied in several subclasses of regular languages other than ideals: prefix-, suffix-, bifix-, and factor-free languages [8, 12]; star-free languages [7, 10]; R- and J-trivial languages [6]; finite/cofinite and reverse definite languages [7]. This problem can be quite challenging, depending on the subclass; in the present case it is easy for right ideals but much more difficult for left- and two-sided ideals (defined below).

As syntactic complexity bounds the maximal size of the transition semigroup, it provides a natural bound on the time and space complexities of algorithms dealing with transition semigroups. For example, a simple algorithm determining whether the language of a given minimal DFA is star-free [23] requires the enumeration of all transformations and checking whether they do not contain non-trivial cycles. A language is star-free if it can be generated from finite languages by using only Boolean operations and product (concatenation), but not star; equivalently, its syntactic semigroup is group-free, that is, has no non-trivial subgroups.

Maximal transition semigroups also play an important role in the study of *most complex* languages [3] belonging to a given subclass. These are languages that meet all the upper bounds on the state complexities of Boolean operations, product, star, and reversal, have maximal syntactic semigroups and most complex atoms [13].

In contrast to the *syntactic monoid* of the language, the syntactic semigroup may or may not contain the neutral element (the identity transformation). The presence of letters acting as identity is often important in the case of state complexity of binary operations. Moreover, the syntactic semigroup is more suitable to characterize some classes of languages, which have a description in terms of semigroups. For example, in the class of (co)finite languages all transformations must admit a certain linear order of the states [15], and the identity transformation cannot be present; the latter condition would not be distinguished by the syntactic monoid.

In this paper we study the syntactic complexities of *right ideals* (satisfying the equation $L = L\Sigma^*$), *left ideals* (satisfying $L = \Sigma^*L$), and *two-sided ideals* (satisfying $L = \Sigma^*L\Sigma^*$). Ideals are fundamental objects in semigroup theory. They appear in the theoretical computer science literature in 1965 [26] and continue to be of interest. Ideal languages are special cases of *convex languages* (see e.g. [9]), and they are complements of prefix-, suffix-, factor-, and subword-closed languages. Besides being of theoretical interest, ideals also play a role in algorithms for pattern matching. For this application, a *text* is represented by a word w over some alphabet Σ. A *pattern* is a language L over Σ. An occurrence of a pattern represented by L in text w is a triple (u, x, v) such that $w = uxv$ and x is in L. Searching text w for words in L is equivalent to looking for prefixes of w that belong to the language Σ^*L, which is the left ideal generated by L, or looking for factors of w that belong to $\Sigma^*L\Sigma^*$ [16].

The state complexity of operations on the classes of ideal languages was studied by Brzozowski, Jirásková and Li [4]. The same problem for the classes of prefix-, suffix-, factor-, and subword-closed languages was studied by Han and K. Salomaa [17], Han, K. Salomaa, and Wood [18], and Brzozowski, Jirásková and Zou [5].

We refer the reader to these papers for a discussion of past work on this topic and additional references.

The set of all n^n transformations of a set Q_n of n elements is a monoid under composition of transformations, with identity as the unit element. In 1970, Maslov [22] dealt with the generators of the semigroup of all transformations in the setting of finite automata. Holzer and König [19], and independently Krawetz, Lawrence, and Shallit [20] studied the syntactic complexity of unary and binary regular languages. Recently, syntactic complexity has been studied in several subclasses of regular languages other than ideals: prefix-, suffix-, bifix-, and factor-free languages [8, 12]; star-free languages [7, 10]; R- and J-trivial languages [6].

We define our terminology and notation in Section 2, and give some basic properties of syntactic complexity in Section 3. The syntactic complexities of right, left, and two-sided ideals are treated in Sections 4–6, and Section 7 concludes the paper. As mentioned above, closed languages are complements of ideal languages. Since syntactic complexity is preserved under complementation, our proofs are for ideals only. The syntactic complexity of all-sided ideals remains open.

In the proof for the upper bounds for left and two-sided ideals we use the method of injective function, which is generally applicable for other subclasses of regular languages (see [12] for suffix-free and [31] for bifix-free languages). The proofs presented here are the first that apply this method to syntactic complexity.

A part of the results in this paper previously appeared in conference proceedings: In 2011 in [14] syntactic complexity of right ideals was established and lower bounds for the classes of left and two-sided ideals were presented. In 2014 in [11] incomplete proofs of the upper bounds for syntactic complexity of left and two-sided ideals were presented.

2 Preliminaries

If Σ is an alphabet (a non-empty finite set), then Σ^* is the free monoid generated by Σ, and Σ^+ is the free semigroup generated by Σ. A *word* is any element of Σ^*, and the empty word is ε. The length of a word $w \in \Sigma^*$ is $|w|$. A *language* over Σ is any subset of Σ^*.

If $w = uxv$ for some $u, v, x \in \Sigma^*$, then u is a *prefix* of w, v is a *suffix* of w, and x is a *factor* of w. A prefix or suffix of w is also a factor of w. If $w = u_1v_1u_2v_2 \cdots u_kv_ku_{k+1}$, where the u_i and v_i are in Σ^*, then $v_1v_2 \cdots v_k$ is a *subword* of w. A language L is *prefix-closed* if $w \in L$ implies that every prefix of w is also in L. In an analogous way, we define *suffix-closed*, *factor-closed*, and *subword-closed*. We refer to all four types as *closed languages*.

The *shuffle* $u \sqcup\!\sqcup v$ *of two words* $u, v \in \Sigma^*$ is defined as follows:

$$u \sqcup\!\sqcup v = \{u_1v_1 \cdots u_kv_k \mid u = u_1 \cdots u_k, v = v_1 \cdots v_k, u_1, \ldots, u_k, v_1, \ldots, v_k \in \Sigma^*\}.$$

The *shuffle of two languages* K and L is defined by

$$K \sqcup\!\sqcup L = \bigcup_{u \in K, v \in L} u \sqcup\!\sqcup v.$$

A language $L \subseteq \Sigma^*$ is a *right ideal* (respectively, *left ideal*, *two-sided ideal*, *all-sided ideal*) if it is non-empty and satisfies $L = L\Sigma^*$ (respectively, $L = \Sigma^* L$, $L = \Sigma^* L \Sigma^*$, $L = \Sigma^* \sqcup\!\sqcup L$). We refer to all four of these types as *ideal languages* or simply *ideals*.

Proposition 1 *Suppose L is a language over Σ^* and $L \neq \Sigma^*$. Let $\overline{L} = \Sigma^* \setminus L$ be the complement of L. Then the following hold:*

- *L is prefix-closed if and only if $\overline{L}$ is a right ideal.*
- *L is suffix-closed if and only if $\overline{L}$ is a left ideal.*
- *L is factor-closed if and only if $\overline{L}$ is a two-sided ideal.*

Proof The claim for factor-closed languages was proved in [21]. The proof for prefix-closed languages [1] parallels the proof in [21], and that for suffix-closed languages follows by the dual argument. □

A *transformation* of a set Q_n of n elements is a mapping of Q_n *into* itself, whereas a *permutation* of Q_n is a mapping of Q_n *onto* itself. In this paper we consider only transformations of finite sets, and we assume without loss of generality that $Q_n = \{0, 1, \dots, n-1\}$. An arbitrary transformation has the form

$$t = \begin{pmatrix} 0 & 1 & \cdots & n-2 & n-1 \\ q_0 & q_1 & \cdots & q_{n-2} & q_{n-1} \end{pmatrix},$$

where $q_k \in Q_n$ for $0 \le k \le n-1$. The image of element q under transformation t is denoted by qt. The *identity* transformation $\mathbf{1}$ maps each element to itself. For $k \ge 2$, a transformation (permutation) s of a set $P = \{p_0, p_1, \dots, p_{k-1}\} \subseteq Q_n$ is a *k-cycle* if $p_0 s = p_1, p_1 s = p_2, \dots, p_{k-2} s = p_{k-1}, p_{k-1} s = p_0$. If a transformation t on Q_n acts on $P \subseteq Q_n$ like a k-cycle then t is said to *have a k-cycle*. A k-cycle is denoted by $(p_0, p_1, \dots, p_{k-1})$ when it is viewed as a transformation of P. If t is a transformation of Q_n, has a k-cycle $(p_0, p_1, \dots, p_{k-1})$ of P, and acts as identity on $Q_n \setminus P$, then we denote t also by $(p_0, p_1, \dots, p_{k-1})$. A 2-cycle (p_0, p_1) is called a *transposition*. A transformation is *constant* if it maps all states to a single state q; it is denoted by $(Q \to q)$. A transformation that maps a single state p to q and keeps $Q \setminus \{p\}$ unchanged is denoted by $(p \to q)$. A transformation mapping p to q_p for $p = 0, \dots, n-1$ is sometimes denoted by $[q_0, \dots, q_{n-1}]$.

The following facts are well-known [28, 30]:

Proposition 2 *The complete transformation monoid $\mathcal{T}_n$ of size n^n can be generated by any cyclic permutation of n elements together with a transposition of any two elements adjacent in the cyclic permutation, and a singular (non-invertible) transformation of rank (image size) $n-1$. In particular, $\mathcal{T}_n$ can be generated by* $(0, 1, \dots, n-1)$, $(0, 1)$ *and* $(n-1 \to 0)$. *Moreover, $\mathcal{T}_n$ cannot be generated by fewer than three generators for $n \ge 3$.*

Remark 1 Let T_n' be a transformation semigroup that requires at least g generators. Suppose T_n contains T_n' as a subsemigroup. If for every $t \in T_n'$, no transformation from $T_n \setminus T_n'$ can be used to generate t, then any set of generators of T_n contains at least g generators from T_n'.

Proof Let G be a set of generators of T_n. Let $t \in T_n'$. Since $t \in T_n$, it is generated by G. Since generators from $T_n \setminus T_n'$ cannot be used, t is generated by generators from $G \cap T_n'$. Thus $G \cap T_n'$ generates T_n', and so G contains at least g generators. □

An equivalence relation $\sim$ on Σ^* is a *right congruence* if, for all $x, y \in \Sigma^*$, $x \sim y \Leftrightarrow xv \sim yv$, for all $v \in \Sigma^*$. It is a *congruence* if $x \sim y \Leftrightarrow uxv \sim uyv$, for all $u, v \in \Sigma^*$.

For any language $L \subseteq \Sigma^*$, define the *Nerode right congruence* [25] $\sim_L$ of L by

$$x \sim_L y \text{ if and only if } xv \in L \Leftrightarrow yv \in L, \text{ for all } v \in \Sigma^*. \tag{1}$$

The *left quotient*, or simply *quotient*, of a language L by a word w is the language $w^{-1}L = \{x \in \Sigma^* \mid wx \in L\}$. Evidently, $x^{-1}L = y^{-1}L$ if and only if $x \sim_L y$. Thus, each equivalence class of this right congruence corresponds to a distinct quotient of L. Let $K = \{K_0, \dots, K_{n-1}\}$ be the set of quotients of a regular language L; by convention, we let $K_0 = L = \varepsilon^{-1}L$. The number of distinct quotients of L is the *quotient complexity* $\kappa(L)$ of L.

The *Myhill congruence* [24] $\approx_L$ of L is defined by

$$x \approx_L y \text{ if and only if } uxv \in L \Leftrightarrow uyv \in L \text{ for all } u, v \in \Sigma^*. \tag{2}$$

This congruence is also known as the *syntactic congruence* of L. The semigroup $\Sigma^+/{\approx_L}$ of equivalence classes of the relation $\approx_L$ is the *syntactic semigroup* of L, and $\Sigma^*/{\approx_L}$ is the *syntactic monoid* of L. The *syntactic complexity* $\sigma(L)$ of L is the cardinality of its syntactic semigroup.

A *deterministic finite automaton (DFA)* is a quintuple $\mathcal{D} = (Q, \Sigma, \delta, q_0, F)$, where Q is a finite, non-empty set of *states*, Σ is an *alphabet*, $\delta\colon Q \times \Sigma \to Q$ is the *transition function*, $q_0 \in Q$ is the *initial state*, and $F \subseteq Q$ is the set of *final states*. As usual, δ is extended to a function from $Q \times \Sigma^*$ to Q. By the *language of a state* q of $\mathcal{D}$ we mean the language K_q accepted by the automaton $(Q, \Sigma, \delta, q, F)$. States p and q are *equivalent* if $K_p = K_q$. A state q is *reachable* if $\delta(q_0, w) = q$ for some $w \in \Sigma^*$. A DFA is *minimal* if every state is reachable and no two states are equivalent. This implies that the number of states of a minimal DFA is minimal.

Each word w of Σ^* induces a transformation t as follows: $qt = \delta(q, w)$ for all $q \in Q$. The fact that w induces transformation t is denoted by $w\colon t$. The *transition semigroup* of a DFA is the set of transformations $q \mapsto \delta(q, w)$ for all $q \in Q$, $w \in \Sigma^+$ induced by words of Σ^+ on the set of states. The transition semigroup of the quotient DFA of L is isomorphic to the syntactic semigroup of L [29].

The *quotient automaton* of L is $\mathcal{D} = (K, \Sigma, \delta, L, F)$, where $\delta(K_q, a) = a^{-1}K_q$, and $F = \{K_q \mid \varepsilon \in K_q\}$. Since the number of distinct quotients of L is precisely the number of states in the quotient automaton, the quotient automaton is always minimal, and so quotient complexity is the same as state complexity.

3 Syntactic Complexity of Languages with Special Quotients

We now present some basic properties of syntactic complexity.

Proposition 3 *For any $L \subseteq \Sigma^*$ with $\kappa(L) = n > 1$, we have $n - 1 \leq \sigma(L) \leq n^n$.*

Proof Let $\mathcal{D} = (K, \Sigma, \delta, L, F)$ be the quotient automaton of L. Since every state other than L has to be reachable from the initial state L by a non-empty word, there must be at least $n - 1$ transformations. If $\Sigma = \{a\}$ and $L = a^{n-1}a^*$, then $\kappa(L) = n$, and $\sigma(L) = n - 1$; so the lower bound $n - 1$ is achievable. The upper bound is n^n, and by Proposition 2 this upper bound is achievable if $|\Sigma| \geq 3$. The upper bound is reachable with $|\Sigma| = 2$ for $n = 2$ by the language $(b \cup aa \cup ab)^*$, and with $|\Sigma| = 1$ for $n = 1$ by the language Σ^*. □

If one of the quotients of L is $\emptyset$ (respectively, $\{\varepsilon\}$, Σ^*, Σ^+), then we say that *L has $\emptyset$* (respectively, $\{\varepsilon\}$, Σ^*, Σ^+). A quotient $w^{-1}L$ of a language L is *uniquely reachable* [2] if $x^{-1}L = w^{-1}L$ implies that $x = w$. If $(wa)^{-1}L$ is uniquely reachable for $a \in \Sigma$, then so is $w^{-1}L$. Thus, if L has a uniquely reachable quotient, then L itself is uniquely reachable by ε, i.e., a minimal automaton of L is *non-returning* [17].

Theorem 1 (**Special Quotients**) *Let $L \subseteq \Sigma^*$ and let $\kappa(L) = n \geq 1$.*

1. *If L has $\emptyset$ or Σ^*, then $\sigma(L) \leq n^{n-1}$.*
2. *If L has $\{\varepsilon\}$ or Σ^+, then $\sigma(L) \leq n^{n-2}$.*
3. *If L is uniquely reachable, then $\sigma(L) \leq (n-1)^n$.*
4. *If $w^{-1}L$ is uniquely reachable by $w \in \Sigma^*$ with $0 \leq |w| \leq n - 1$, then $\sigma(L) \leq |w| + (n - 1 - |w|)^n$.*

Moreover, all the bounds shown in Table 1 *hold.*

Table 1 Upper bounds on syntactic complexity for languages with special quotients

$\emptyset$	Σ^*	$\{\varepsilon\}$	Σ^+	$\sigma(L) \leq$	if also L is ur	if also $a^{-1}L$ is ur
$\surd$				n^{n-1}	$(n-1)^{n-1}$	$1 + (n-3)^{n-1}$
	$\surd$			n^{n-1}	$(n-1)^{n-1}$	$1 + (n-3)^{n-1}$
$\surd$		$\surd$		n^{n-2}	$(n-1)^{n-2}$	$1 + (n-4)^{n-2}$
	$\surd$		$\surd$	n^{n-2}	$(n-1)^{n-2}$	$1 + (n-4)^{n-2}$
$\surd$	$\surd$			n^{n-2}	$(n-1)^{n-2}$	$1 + (n-4)^{n-2}$
$\surd$	$\surd$		$\surd$	n^{n-3}	$(n-1)^{n-3}$	$1 + (n-5)^{n-3}$
$\surd$	$\surd$	$\surd$		n^{n-3}	$(n-1)^{n-3}$	$1 + (n-5)^{n-3}$
$\surd$	$\surd$	$\surd$	$\surd$	n^{n-4}	$(n-1)^{n-4}$	$1 + (n-6)^{n-4}$

The abbreviation "ur" stands for "uniquely reachable". The a in the last column is in Σ

Proof Suppose that $L \subseteq \Sigma^*$, $n \geq 1$, and $\kappa(L) = n$.

1. Since $a^{-1}\emptyset = \emptyset$ for all $a \in \Sigma$, there are only $n-1$ states in the quotient automaton with which one can distinguish two transformations. Hence there are at most n^{n-1} transformations. If L has Σ^*, then $a^{-1}\Sigma^* = \Sigma^*$, for all $a \in \Sigma$, and the same argument applies.
2. Since $a^{-1}\{\varepsilon\} = \emptyset$ for all $a \in \Sigma$, a language L has $\emptyset$ if L has $\{\varepsilon\}$. Now there are two states that do not contribute to distinguishing among different transformations. Dually, $a^{-1}\Sigma^+ = \Sigma^*$ for all $a \in \Sigma$, and the same argument applies.
3. If L is uniquely reachable then $w^{-1}L = L$ implies $w = \varepsilon$. Thus L does not appear in the image of any transformation by a word in Σ^+, and there remain only $n-1$ choices for each of the n states.
4. If $w^{-1}L$ is uniquely reachable, then so is $x^{-1}L$ for every prefix x of w. Hence for each prefix x of w, $x^{-1}L$ appears only in one transformation, and there are $|w|$ such transformations. All the other transformations map every quotient $x^{-1}L$ to $y^{-1}L$, where y is not a prefix of w. Therefore there can be at most $(n-1-|w|)^n$ other transformations.

The remaining entries in Table 1 are easily verified: every transformation fixes $\emptyset$, Σ^*, maps $\{\varepsilon\}$ to $\emptyset$, and maps Σ^+ to Σ^*, so these quotients are removed from counting possible mappings for a quotient. □

4 Right Ideals and Prefix-Closed Languages

In this section we prove that the syntactic complexity of right ideals is n^{n-1}. First we define a witness DFA that meets this bound.

Definition 1 (**Witness: Right Ideals**) For $n \geq 3$, let $\mathcal{W}_n = (Q_n, \Sigma, \delta_{\mathcal{W}}, 0, \{n-1\})$, be the DFA in which $\Sigma = \{a, b, c, d\}$, $a\colon (0, \dots, n-2)$, $b\colon (0, 1)$, $c\colon (n-2 \to 0)$, and $d\colon (n-2 \to n-1)$. For $n = 3$ inputs a and b induce the same transformation; hence $\Sigma = \{a, c, d\}$ suffices. Furthermore, let $\mathcal{W}_2 = (Q_2, \{a, b\}, \delta_{\mathcal{W}}, 0, \{1\})$, where $a\colon (0 \to 1)$, and $b\colon \mathbf{1}$, and let $\mathcal{W}_1 = (Q_1, \{a\}, \delta_{\mathcal{W}}, 0, \{0\})$, where $a\colon \mathbf{1}$. Let $L_n = L(\mathcal{W}_n)$.

The structure of the DFA of Definition 1 is shown in Fig. 2 for $n \geq 3$.

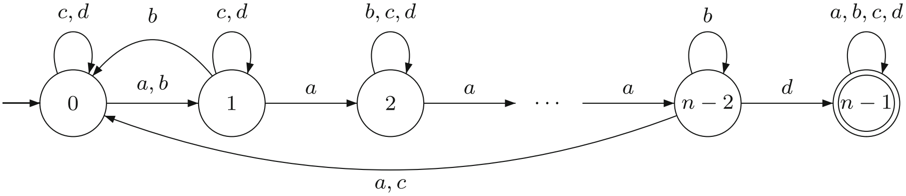

Fig. 2 Quotient DFA $\mathcal{W}_n$ of a right ideal with n^{n-1} transformations

Let $\mathbf{W}_{\mathrm{ri}}$ be the transition semigroup of the witness $\mathcal{W}_n$.

Lemma 1 *The DFA $\mathcal{W}_n$ of Definition 1 is minimal, accepts a right ideal, and its transition semigroup $\mathbf{W}_{\mathrm{ri}}$ has size n^{n-1}.*

Proof If $n \le 2$ this is easily verified; here $L_1 = \Sigma^*$ and $L_2 = \Sigma^* a \Sigma^*$.

For $n \ge 3$, any state q with $0 \le q \le n-2$ is non-final, accepts $a^{n-2-q}d$, and no other such state accepts this word. Since $n-1$ is final, all states are distinguishable. Since $\mathcal{W}_n$ has exactly one final state and that state accepts Σ^*, L_n is a right ideal.

For the syntactic complexity, observe that inputs a, b, and c restricted to Q_{n-1} can induce any transformation of Q_{n-1} (Proposition 2); hence all $(n-1)^{n-1}$ transformations that fix $n-1$ can be performed by $\mathcal{W}_n$. Also observe that any transformation $(q \to n-1)$ for $q \in \{0, \ldots, n-3\}$ is induced by $a^{n-2-q}da^{q+1}$.

Note that every transformation from the transition semigroup $\mathbf{W}_{\mathrm{ri}}$ fixes state $n-1$. Let t be any transformation such that $(n-1)t = n-1$. There are n^{n-1} such transformations, and we will show that all of them are generated. Let $\{p_1, \ldots, p_k\}$ be the set of all states from $Q \setminus \{n-1\}$ that are mapped by t to $n-1$. Then t can be generated by $(p_1 \to n-1) \cdots (p_k \to n-1)t'$, where t' fixes $n-1$ and all states p_i, and acts as t on the other states; thus it is a transformation of Q_{n-1} if restricted to Q_{n-1} and can be generated by a, b, and c. □

We are now in a position to state our main theorem of this section.

Theorem 2 (**Right Ideals and Prefix-Closed Languages**) *Suppose that $L \subseteq \Sigma^*$ and $\kappa(L) = n$. If L is a right ideal or a prefix-closed language, then $\sigma(L) \le n^{n-1}$. This bound is tight for $n = 1$ if $|\Sigma| \ge 1$, for $n = 2$ if $|\Sigma| \ge 2$, for $n = 3$ if $|\Sigma| \ge 3$, and for $n \ge 4$ if $|\Sigma| \ge 4$. Moreover, the sizes of the alphabet cannot be reduced.*

Proof For $n \ge 4$, every transformation in the transition semigroup of a minimal DFA of any right ideal with n quotients must fix state $n-1$; hence the size of this semigroup cannot exceed n^{n-1}. By Lemma 1 this bound is tight.

It is easy to verify that the alphabet cannot be smaller if $n \le 3$. Let $n \ge 4$. The set of transformations in the largest transition semigroup must contain every transformation t that maps Q_{n-1} to Q_{n-1} and fixes $n-1$; otherwise, the bound cannot be met. Thus, none of the generators of this semigroup can map a state from Q_{n-1} to $n-1$. When restricted to Q_{n-1}, the transformations in this semigroup must form the full transformation semigroup of Q_{n-1} with $n-1 \ge 3$ states. So by Remark 1, from Proposition 2 we know that there must be at least three generators of these transformations, say a, b, c. As noted above, none of $\{a, b, c\}$, extended to Q_n by adding the mapping of $n-1$ to $n-1$, can map a state from Q_{n-1} to $n-1$. So we need at least one more generator, say d, which maps a state from Q_{n-1} to $n-1$. Altogether, at least four generators are needed.

Since prefix-closed languages are complements of right ideals and the syntactic complexity is preserved by complementation, the result is the same for prefix-closed languages. □

Remark 2 A maximal transition semigroup of the quotient DFA of a right ideal contains all transformations of Q_n that fix state $n-1$. Hence there is only one maximal transition semigroup for right ideals, which is $\mathbf{W}_{\text{ri}}$.

5 Left Ideals and Suffix-Closed Languages

5.1 Basic Properties

Let $\mathcal{D}_n = (Q_n, \Sigma_{\mathcal{D}}, \delta_{\mathcal{D}}, 0, F)$ be a minimal DFA, and let T_n be its transition semigroup. Consider the sequence $(0, 0t, 0t^2, \dots)$ of states obtained by applying a transformation $t \in T_n$ repeatedly, starting with the initial state. Since Q_n is finite, there must eventually be a repeated state, that is, there must exist i and j such that $0, 0t, \dots, 0t^i, 0t^{i+1}, \dots, 0t^{j-1}$ are distinct, but $0t^j = 0t^i$; the integer $j-i$ is the *period* of t. If the period is 1, t is said to be *initially aperiodic*. If t is initially aperiodic, then its sequence is $0, 0t, \dots, 0t^{j-1} = 0t^j$.

Lemma 2 *If $\mathcal{D}_n$ is the quotient DFA of a left ideal, all the transformations in T_n are initially aperiodic, and the empty set is not a quotient of L.*

Proof Let t be a transformation that is not initially aperiodic. Then there exist i, j such that $p_i = 0t^i = 0t^j = p_j$ for some $i < j$, where $j - i \geq 2$. Let w be a word that induces t. Since $\mathcal{D}_n$ is minimal, states p_i and p_{j-1} must be distinguishable, say by word $x \in \Sigma^*$. If $w^i x \in L$, then $w^{j-1}x = w^i w^{j-i-1} x = w^{j-i-1}(w^i x) \notin L$, contradicting the assumption that L is a left ideal. If $w^{j-1}x \in L$, then $w^j x = w(w^{j-1}x) \notin L$, again contradicting that L is a left ideal.

For the second claim, we know that a left ideal is non-empty by definition. So suppose that $w \in L$. If L has the empty quotient, say $x^{-1}L = \emptyset$, then $xw \notin L$, which contradicts the assumption that L is a left ideal. □

Example 1 Note that the conditions of Lemma 2 are not sufficient. For $\Sigma = \{a, b\}$, the language $L = b \cup \Sigma^* a$ satisfies the conditions, but is not a left ideal because $b \in L$ but $ab \notin L$. Its quotient automaton is shown in Fig. 3.

If the final state is 2 instead of 1, the language becomes $L' = \Sigma\Sigma^* b = \Sigma^*\Sigma b$, which *is* a left ideal. The languages L and L' have the same syntactic semigroup, but one is a left ideal while the other is not.

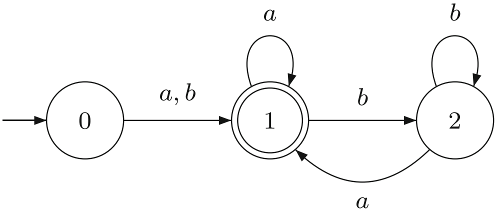

Fig. 3 Quotient DFA of a language that is not a left ideal

The following remark was proved in [4]:

Remark 3 A language $L \subseteq \Sigma^*$ is a left ideal if and only if for all $x, y \in \Sigma^*$, $y^{-1}L \subseteq (xy)^{-1}L$. Hence, if $x^{-1}L \neq L$, then $L \subset x^{-1}L$ for any $x \in \Sigma^+$.

Proof If L is a left ideal then for all $x, y, w \in \Sigma^*$, we have $yw \in L$ implies $xyw \in L$, that is, $w \in y^{-1}L$ implies $w \in (xy)^{-1}L$.

For the other direction, if for some $x, y \in \Sigma^*$ there is $w \in y^{-1}L$ such that $w \notin (xy)^{-1}L$, then $x \neq \varepsilon$ and $yw \in L$ but $xyw \notin L$, which contradicts that L is a left ideal. □

It is useful to restate this observation it terms of the states of $\mathcal{D}_n$. For DFA $\mathcal{D}_n$ and states $p, q \in Q_n$, we write $p \prec q$ if $K_p \subset K_q$. Also, we write $p \preceq q$ if $K_p \subseteq K_q$.

Remark 4 A DFA $\mathcal{D}_n$ is a minimal DFA of a left ideal if and only if for all $s, t \in T_n \cup \{\mathbf{1}\}$, $0t \preceq 0st$. Equivalently, since $q = 0s$ for some s, for every $q \in Q_n \setminus \{0\}$ we have $0 \prec q$.

Remark 5 In a minimal DFA $\mathcal{D}_n$ of a left ideal, if $r \in Q_n$ has a t-predecessor, that is, if there exists $q \in Q_n$ such that $qt = r$, then $0t \preceq r$. In particular, if r appears in a cycle of t or is a fixed point of t, then $0t \preceq r$.

Proof This follows because $0 \preceq q$ and so $0t \preceq qt = r$ by Remark 4. □

We consider chains of the form $K_{i_1} \subset K_{i_2} \subset \cdots \subset K_{i_h}$, where the K_{i_j} are quotients of L. If L is a left ideal, the smallest element of any maximal-length chain is always L. Alternatively, we consider chains of states starting from 0 and strictly ordered by $\prec$.

Proposition 4 *For $t \in T_n$ and $p, q \in Q_n$, $p \prec q$ implies $pt \preceq qt$. If $p \prec pt$, then $p \prec pt \prec \cdots \prec pt^k = pt^{k+1}$ for some $k \geq 1$. Similarly, $p \succ q$ implies $pt \succeq qt$, and $p \succ pt$ implies $p \succ pt \succ \cdots \succ pt^k = pt^{k+1}$ for some $k \geq 1$.*

Proof Since $\subseteq$ is a partial order on quotients, by definition of $\prec$, if $K_p \subset K_q$ then $w^{-1}K_p \subseteq w^{-1}K_q$, where w is a word inducing t. This applied iteratively yields $p \prec pt \prec \cdots \prec pt^k = pt^{k+1}$ for some $k \geq 1$, because there are finitely many quotients ($k \leq n$). The same hold dually for $\succ$. □

5.2 Lower Bound

We now show that the syntactic complexity of the following DFA of a left ideal is $n^{n-1} + n - 1$.

Definition 2 (Witness: Left Ideals) For $n \geq 3$, let $\mathcal{W}_n = (Q_n, \Sigma_{\mathcal{W}}, \delta_{\mathcal{W}}, 0, \{n-1\})$, be the DFA in which $\Sigma_{\mathcal{W}} = \{a, b, c, d, e\}$, $a\colon (1, \ldots, n-1)$, $b\colon (1, 2)$, $c\colon (n-1 \to 1)$, $d\colon (n-1 \to 0)$, and $e\colon (Q_n \to 1)$. For $n = 3$, a and b coincide, and we can use $\Sigma_{\mathcal{W}} = \{a, c, d, e\}$. Also, let $\mathcal{W}_2 = (Q_2, \{a, b, c\}, \delta_{\mathcal{W}}, 0, \{1\})$, where $a\colon (0 \to 1)$,

$b\colon \mathbf{1}$, and $c\colon (Q_2 \to 1)$, and let $\mathcal{W}_1 = (Q_1, \{a\}, \delta, 0, \{0\})$, where $a\colon \mathbf{1}$. Let $L_n = L(\mathcal{W}_n)$.

The structure of the DFA of Definition 2 is shown in Fig. 4 for $n \geq 3$.

Lemma 3 *The DFA of Definition 2 is minimal, accepts a left ideal, and has transition semigroup of size $n^{n-1} + n - 1$ that contains all the transformations fixing 0 and all the constant transformations.*

Proof State 0 does not accept a^i for any i, whereas state i with $1 \leq i \leq n-2$ accepts a^{n-1-i}, and no other state j with $1 \leq j \leq n-2$ accepts this word. Since $n-1$ is the only final state, all states are distinguishable.

To prove that L is a left ideal it suffices to show that for any $w \in L$, we also have $xw \in L$ for every $x \in \Sigma$. This is obvious if $x \in \Sigma \setminus \{e\}$. If $w \in L$, then w has the form $w = uev$, where $\delta_{\mathcal{W}}(0, u) = 0$, $\delta_{\mathcal{W}}(0, ue) = 1$, and v is accepted from state 1. But $\delta_{\mathcal{W}}(0, eue) = 1$, and since v is accepted from 1, we have $euev = ew \in L_n$. Thus L_n is a left ideal.

In $\mathcal{W}_n$, the transformations induced by a, b, and c restricted to $Q_n \setminus \{0\}$ generate all the transformations of the last $n-1$ states (Proposition 2). Together with the transformation of d, they generate all transformations of Q_n that fix 0, and the number of such transformations is n^{n-1}. To see this, consider any transformation t that fixes 0. If some states from $\{1, \dots, n-1\}$ are mapped to 0 by t, we can map them first to $n-1$ and $n-1$ to one of them by the transformations of a, b, and c, and then map $n-1$ to 0 by the transformation of d.

Also the words of the form ea^i for $i \in \{0, \dots, n-2\}$ induce constant transformations $(Q_n \to i+1)$. Hence the transition semigroup of $\mathcal{W}_n$ contains all the constant transformations of Q_n (where $(Q_n \to 0)$ has been already counted). Altogether, there are $n^{n-1} + n - 1$ transformations in the transition semigroup of $\mathcal{W}_n$. □

Example 2 The maximal-length chains of quotients in $\mathcal{W}_n$ have length 2. However, in other left ideals maximal-length chains can be as long as n. For this let $n \geq 2$,

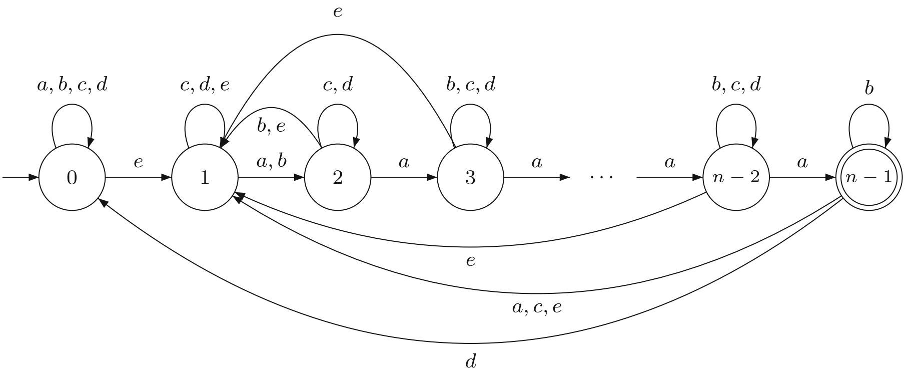

Fig. 4 Quotient DFA $\mathcal{W}_n$ of a left ideal with $n^{n-1} + n - 1$ transformations

$\Sigma = \{a, b\}$ and $L = \Sigma^* a^{n-1}$; then L has n quotients and a maximal-length chain of length n.

Proof A maximal-length chain always starts at 0; suppose it ends with q. If there is a $p \in Q_n \setminus \{0, q\}$ such that $p \prec q$, then $K_p \subset K_q$, which contradicts that $a^{n-1-p} \in K_p$ and $a^{n-1-p} \notin K_q$.

In $L = \Sigma^* a^{n-1}$, we have the unique maximal-length chain consisting of all quotients:

$$\Sigma^* a^{n-1} \subset \Sigma^* a^{n-2} \subset \cdots \subset \Sigma^*. \qquad \square$$

We will see that the maximal length of chains of quotients is an important structural feature; in particular, to meet the bound for syntactic complexity by both left and two-sided ideals, the maximal length of the chains must be the smallest possible.

5.3 Upper Bound

The derivation of the upper bound $n^{n-1} + n - 1$ for left ideals is much more difficult that for right ideals. We begin with the easy cases where $n \in \{1, 2\}$.

Remark 6 If $n = 1$, the only left ideal is Σ^* and the transition semigroup of its minimal DFA satisfies the bound $1^0 + 1 - 1 = 1$. If $n = 2$, there are only three allowed transformations, since the transposition (0, 1) is not initially aperiodic and is ruled out by Lemma 2. Thus the bound $2^1 + 2 - 1 = 3$ holds.

Let $\mathcal{D}_n = (Q_n, \Sigma_{\mathcal{D}}, \delta_{\mathcal{D}}, 0, F)$ be a minimal DFA of an arbitrary left ideal with n quotients and let T_n be the transition semigroup of $\mathcal{D}_n$. Let $\mathbf{W}_{\text{li}}$ be the transition semigroup of the witness DFA $\mathcal{W}_n$ of Definition 2.

Lemma 4 *If $n \geq 3$ and a maximal-length chain in $\mathcal{D}_n$ strictly ordered by $\prec$ has length 2, then $|T_n| \leq n^{n-1} + n - 1$ and T_n is a subsemigroup of* $\mathbf{W}_{\text{li}}$.

Proof Consider an arbitrary transformation $t \in T_n$ and let $p = 0t$. If $p = 0$, then any state other than 0 can possibly be mapped by t to any one of the n states; hence there are at most n^{n-1} such transformations. All of these transformations are in $\mathbf{W}_{\text{li}}$ by the proof of Lemma 3.

If $p \neq 0$, then $0 \prec p$. Consider any state $q \notin \{0, p\}$; by Remark 4, $0t = p \preceq qt$. If $p \neq qt$, then $p \prec qt$. But then we have the chain $0 \prec p \prec qt$ of length 3, contradicting our assumption. Hence we must have $p = qt$, and so t is the constant transformation $t = (Q_n \to p)$. Since p can be any one of the $n - 1$ states other than 0, we have at most $n - 1$ such transformations. Since all of these transformations are in $\mathbf{W}_{\text{li}}$ by Lemma 3, T_n is a subsemigroup of $\mathbf{W}_{\text{li}}$. $\square$

Lemma 5 (**Left Ideals, Suffix-Closed Languages**) *If $n \geq 3$ and L is a left ideal or a suffix-closed language with n quotients, then its syntactic complexity is less than or equal to $n^{n-1} + n - 1$.*

Proof Our approach is as follows: We consider a minimal DFA $\mathcal{D}_n = (Q_n, \Sigma_{\mathcal{D}}, \delta_{\mathcal{D}}, 0, F)$ of an arbitrary left ideal with n quotients and let T_n be the transition semigroup of $\mathcal{D}_n$. We also deal with the witness DFA $\mathcal{W}_n = (Q_n, \Sigma_{\mathcal{W}}, \delta_{\mathcal{W}}, 0, \{n-1\})$ of Definition 2 that has the same state set as $\mathcal{D}_n$ and whose transition semigroup is $\mathbf{W}_{\text{li}}$. We will show that there is an injective mapping $f : T_n \to \mathbf{W}_{\text{li}}$, and this will prove that $|T_n| \le |\mathbf{W}_{\text{li}}|$.

It suffices to prove the result for left ideals, since suffix-closed languages are their complements.

In the proof of this lemma we enumerate the following cases illustrated in Fig. 5:

Case 1: $t \in \mathbf{W}_{\text{li}}$.
Case 2: $t \notin \mathbf{W}_{\text{li}}$ and $0t^2 \neq 0t$.
Case 3: $t \notin \mathbf{W}_{\text{li}}$ and $0t^2 = 0t$.

(a): t has a cycle.
(b): t has no cycles and has a fixed point $r \neq p$.
(c): t has no cycles, has no fixed point $r \neq p$, and there is a state r such that $p \prec r$ with $rt = p$.

We now proceed to examine each of these cases.

Case 1: $t \in \mathbf{W}_{\text{li}}$.

Let $f(t) = t$; then obviously f restricted to $\mathbf{W}_{\text{li}}$ is injective.

Case 2: $t \notin \mathbf{W}_{\text{li}}$ and $0t^2 \neq 0t$.

Note that $t \notin \mathbf{W}_{\text{li}}$ implies $0t \neq 0$ by Lemma 3. Let $0t = p$. Since $0t^2 \neq 0t$, we have $p = 0t \prec 0tt = pt$ by Remark 4. Let $p \prec \cdots \prec pt^k = pt^{k+1}$ be the

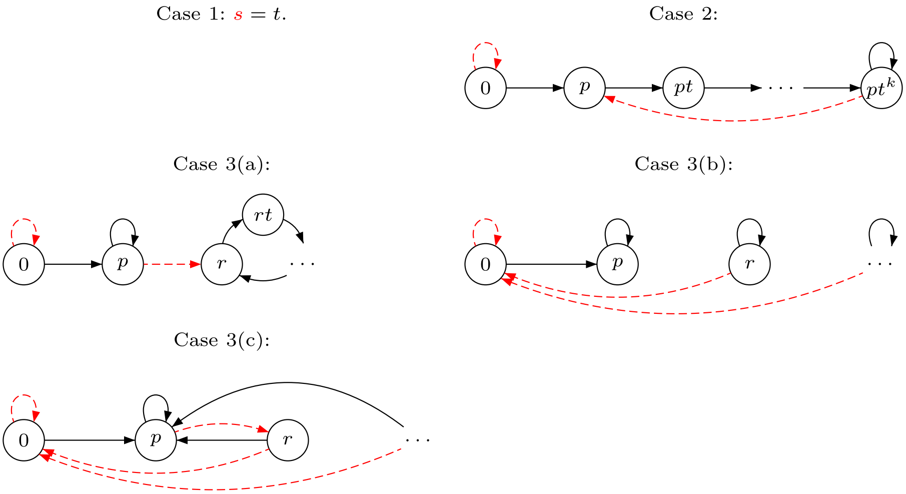

Fig. 5 Map of the cases in the proof of Lemma 5. The transitions of t are represented by *solid lines*, and the modified transitions of s by *dashed red lines*

chain defined from p; this chain is of length at least 2. Let $f(t) = s$, where s is the transformation defined by

$$0s = 0, \quad pt^k s = p, \quad qs = qt \text{ for the other states } q \in Q_n.$$

Transformation s is shown in Fig. 5, Case 2, where the dashed transitions show how s differs from t.

By Lemma 3, $s \in \mathbf{W}_{\mathrm{li}}$. However, $s \notin T_n$, as it contains the cycle $(p, \dots, pt^k)$ with states strictly ordered by $\prec$ in DFA $\mathcal{D}_n$, which contradicts Proposition 4. Since $s \notin T_n$, it is distinct from the transformations defined in Case 1.

In going from t to s, we have added one transition ($0s = 0$) that is a fixed point, and one ($pt^k s = p$) that is not. Since only one non-fixed-point transition has been added, there can be only one cycle in s with states strictly ordered by $\prec$. Since 0 cannot appear in this cycle, p is its smallest element with respect to $\prec$.

Suppose now that $t' \neq t$ is another transformation that satisfies Case 2, that is, $0t' = p' \neq 0$ and $p't' \neq p'$; we will show that $f(t) \neq f(t')$. Define s' for t' as s was defined for t. For a contradiction, assume $s = f(t) = f(t') = s'$.

Like state s, state s' contains only one cycle strictly ordered by $\prec$, and p' is its smallest element. Since we have assumed that $s = s'$, we must have $p = 0t = 0t' = p'$ and the cycles in s and s' must be identical. In particular, $pt^k t = pt^k = p(t')^k t' = p(t')^k$. For q of $Q_n \setminus \{0, pt^k\}$, we have $qt = qs = qs' = qt'$. Hence $t = t'$ – a contradiction. Therefore $t \neq t'$ implies $f(t) \neq f(t')$.

Case 3: $t \notin \mathbf{W}_{\mathrm{li}}$ and $0t^2 = 0t$.

As before, let $0t = p$. Consider any state $q \notin \{0, p\}$; then $0 \prec q$ by Remark 4 and $0t \preceq qt$ by Proposition 4. Thus either $p \prec qt$, or $p = qt$. We consider the following sub-cases:

- **(a):** t has a cycle.

Since t has a cycle, take a state r from the cycle; then r and rt are not comparable under $\preceq$ by Proposition 4, and $p \prec r$ by Remark 5. Let $f(t) = s$, where s is the transformation shown in Fig. 5, Case 3(a), and defined by

$$0s = 0, \quad ps = r, \quad qs = qt \text{ for the other states } q \in Q_n.$$

By Lemma 3, $s \in \mathbf{W}_{\mathrm{li}}$. Suppose that $s \in T_n$; since $p \prec r$, we have $r = ps \preceq rs = rt$ by the definition of s and Proposition 4; this contradicts that r and rt are not comparable. Hence $s \notin T_n$, and so s is distinct from the transformations of Case 1.

We claim that p is not in a cycle of s; this cycle would have to be

$$p \xrightarrow{s} r \xrightarrow{s} rt \xrightarrow{s} \dots \xrightarrow{s} rt^{k-1} \xrightarrow{s} p, \text{ that is, } p \xrightarrow{s} r \xrightarrow{t} rt \xrightarrow{t} \dots \xrightarrow{t} rt^{k-1} \xrightarrow{t} p,$$

for some $k \geq 2$ because $r \neq p = pt$ and $rt \neq p$. Since $p \prec r$ we have $p = pt \prec rt$; but then we have a chain $p \prec rt \prec \cdots \prec rt^k = p$, contradicting Proposition 4.

Since p is not in a cycle of s, it follows that s does not contain a cycle with states strictly ordered by $\prec$, as such a cycle would also be in t. So s is distinct from the transformations of Case 2.

We claim there is a unique state q such that (a) $0 \prec q \prec qs$, (b) $qs \not\preceq qs^2$. First we show that p satisfies these conditions: (a) holds because $ps = r$ and $p \prec r$; (b) holds because $ps = r$, $ps^2 = rt$ and r and rt are not comparable. Now suppose that q satisfies the two conditions, but $q \neq p$. Note that $qs \neq p$, because $qs = p$ implies $qs = p \prec r = qs^2$, contradicting (b). Since $q, qs \notin \{0, p\}$, we have $qt = qs \not\preceq qs^2 = qt^2$. But Proposition 4 for $q \prec qt$ implies that $qt \preceq qt^2$ – a contradiction. Thus p is the only state satisfying these conditions.

If $t' \neq t$ is another transformation satisfying the conditions of this case, we define s' like s. Suppose that $s = f(t) = f(t') = s'$. Since both s and s' contain a unique state p satisfying the two conditions above, we have $0t = 0t' = p$ and $pt = pt' = p$. Since the other states are mapped by s exactly as by t and t', we have $t = t'$.

- **(b):** t has no cycles and has a fixed point $r \neq p$.

Because $0 \prec r$ by Remark 4, $0t \preceq rt$ by Proposition 4. Since r is a fixed point of t, then $p = 0t \preceq rt = r$. Since $r \neq p$, we have $p \prec r$. Let $f(t) = s$, where s is the transformation shown in Fig. 5, Case 3(b), and defined by

$$0s = 0, \quad qs = 0 \text{ for each fixed point } q \notin \{0, p\}, qs = qt \text{ for the other states } q \in Q_n.$$

By Lemma 3, $s \in \mathbf{W}_{\mathrm{li}}$. Suppose that $s \in T_n$; because $p \prec r$, $ps = p$, $rs = 0$, and $ps \preceq rs$ by Proposition 4, we have $p \prec 0$, which is a contradiction. Hence s is not in T_n and so is distinct from the transformations of Case 1. Also, s maps at least one state other than 0 to 0, and so is distinct from the transformations of Case 2 and also from the transformations of Case 3(a).

If $t' \neq t$ is another transformation satisfying the conditions of this case, we define s' like s. Now suppose that $s = f(t) = f(t') = s'$. There is only one fixed point of s other than 0 ($ps = p$), and only one fixed point of s' other than 0 ($p's' = p'$); hence $0t = p = p' = 0t'$. By the definition of s, for each state $q \neq 0$ such that $qs = 0$, we have $qt = q$. Similarly, for each state $q \neq 0$ such that $qs' = 0$, we have $qt' = q$. Hence t and t' agree on these states. Since the remaining states are mapped by s exactly as they are mapped by t and t', we have $t = t'$. Thus we have proved that $t \neq t'$ implies $f(t) \neq f(t')$.

- **(c):** t has no cycles, has no fixed point $r \neq p$, and there is a state r such that $p \prec r$ with $rt = p$.

Let $f(t) = s$, where s is the transformation shown in Fig. 5, Case 3(c), and defined by

$$0s = 0, \quad ps = r, \quad qs = 0 \text{ for each } q \succ p \text{ such that } qt = p,$$

$$qs = qt \text{ for the other states } q \in Q_n.$$

By Lemma 3, $s \in \mathbf{W}_{\mathrm{li}}$. Suppose that $s \in T_n$; because $p \prec r$, $ps = r$, $rs = 0$, and $r = ps \prec rs = 0$ by Proposition 4, we have $r \prec 0$ – a contradiction. Hence $s \notin T_n$ and s is distinct from the transformations of Case 1.

Because s maps at least one state other than 0 to 0 ($rs = 0$), it is distinct from the transformations of Case 2 and 3(a). Also s does not have a fixed point other than 0, while the transformations of Case 3(b) have such a fixed point.

We claim that there is a unique state q such that (a) $0 \prec q \prec qs$ and (b) $qs^2 = 0$. First we show that p satisfies these conditions. By assumption $0 \prec p \prec r$ and $rt = p$; also $rs = 0$ by the definition of s. Condition (a) holds because $0 \prec p \prec r = ps$, and (b) holds because $0 = rs = ps^2$.

Now suppose that $0 \prec q \prec qs$, $qs^2 = 0$ and $q \neq p$. Since $qs \neq 0$, we have $qs = qt$ by the definition of s. Because qt has a t-predecessor, $p \preceq qt$ by Remark 5. Also $qt = qs \neq p$, for $qs = p$ implies $0 = qs^2 = ps = r$ – a contradiction. Hence $p \prec qt$. From $qt = qs$ and $q \prec qs$, we have $q \prec qt$. Since $qs^2 = 0$ we have $(qt)s = 0$ and so $(qt)t = p$, by the definition of s. By Proposition 4, from $q \prec qt$ we have $qt \preceq (qt)t = p$, contradicting $p \prec qt$. So $q = p$.

If $t' \neq t$ is another transformation satisfying the conditions of this case, we define s' like s. Suppose that $s = f(t) = t(t') = s'$. Since s and s' contain a unique state p satisfying the two conditions above, we have $0t = 0t' = p$ and $pt = pt' = p$. Then r and the states $q \succ p$ with $qt = p$ are determined by p, since they are precisely the states $q \succ p$ with $qs = 0$. Since the other states are mapped by s exactly as by t and t', we have $t = t'$, and f is injective restricted to the transformations of this case also.

- **All cases are covered:**

We need to ensure that any transformation t fits in at least one case. It is clear that t fits in Case 1 or 2 or 3. Let $p = 0t$. For Case 3, it is sufficient to show that if (i) $t \notin \mathbf{W}_{\mathrm{li}}$ does not contain a fixed point $r \neq p$, and (ii) there is no state r with $p \prec r$ and $rt = p$, then t contains a cycle and so fits in Subcase 3(c).

First, if there is no r such that $p \prec r$, we claim that t is the constant transformation $(Q_n \to p)$, thus it fits in Case 1. Consider any state $q \in Q_n$ such that $qt \neq p$. Then $p \prec qt$ by Remark 4, contradicting that there is no state $r = qt$ such that $p \prec r$.

So let t be a transformation that fits in Case 3 and satisfies (i) and (ii), and let r be some state such that $p \prec r$. Consider the sequence $r, rt, rt^2, \ldots$. By Remark 5, $p \preceq rt^i$ for all $i \geq 0$. If $rt^k = p$ for some $k \geq 1$, let k be the smallest such number, then $rt^{k-1} \neq p$; we have $p \prec rt^{k-1}$ and $(rt^{k-1})t = p$, contradicting (ii). Since p is the only fixed point by (i), we have $rt^i \neq rt^{i-1}$ for all $i \geq 1$. Since there are finitely many states, $rt^i = rt^j$ for some i and j such that $0 \leq i < j - 1$, and so the states $rt^i, rt^{i+1}, \ldots, rt^j = rt^i$ form a cycle.

We have shown that for every transformation t in T_n there is a corresponding transformation $f(t)$ in $\mathbf{W}_{\mathrm{li}}$, and f is injective. So $|T_n| \leq |\mathbf{W}_{\mathrm{li}}| = n^{n-1} + n - 1$. □

Next we prove that $\mathbf{W}_{\mathrm{li}}$ is the only transition semigroup meeting the bound. It follows that minimal DFAs of left ideals with the maximal syntactic complexity have maximal-length chains of length 2.

Theorem 3 *If T_n has size $n^{n-1} + n - 1$, then $T_n = \mathbf{W}_{\mathrm{li}}$.*

Proof Consider a maximal-length chain of states strictly ordered by $\prec$ in $\mathcal{D}_n$. If its length is 2, then by Lemma 4, T_n is a subsemigroup of $\mathbf{W}_{\mathrm{li}}$. Thus only $T_n = \mathbf{W}_{\mathrm{li}}$ reaches the bound in this case.

Assume now that the length of a maximal-length chain is at least 3. Then there are states p and r such that $0 \prec p \prec r$. Let $R = \{q \mid p \prec q\}$. We show that there exists a transformation s that is in $\mathbf{W}_{\text{li}}$ but not in $f(T_n)$. To define s we use the constant transformation $t = (Q_n \to p)$ as an auxiliary transformation. Note that t fits in Case 3(c) in the proof of Lemma 5 except that $t \in \mathbf{W}_{\text{li}}$. We define s from t according to the rules of Case 3(c):

$0s = 0$, $\quad ps = r$, $\quad qs = 0$ for each $q \in R$,
$qs = qt = p$ for the other states q.

By Lemma 3, $s \in \mathbf{W}_{\text{li}}$.

Let f be the injective function from the proof of Lemma 5. It remains to be shown that there is no transformation $t' \in T_n$ such that $s = f(t')$. The proof that s is different from the transformations $f(t')$ of Cases 1, 2, 3(a) and 3(b) is exactly the same as the corresponding proof in Case 3(c) following the definition of s.

It remains to verify that there is no $t' \in T_n$ in Case 3(c) such that $f(t') = s$. Suppose there is such a t'. Recall that states p and r satisfying $0 \prec p \prec r$ have been fixed by assumption. By the definition of s, state p satisfies the conditions (a) $0 \prec p \prec ps$ and (b) $ps^2 = 0$. We claim that p is the only state satisfying these conditions. Indeed, if $q \neq p$ then either $qs = 0$, $q \not\prec qs = 0$ and (a) is violated, or $qs = p$, $qs^2 = ps = r \neq 0$ and (b) is violated. This observation is used in the proof of Case 3(c) to prove the claim below.

Both t and t' satisfy the conditions of Case 3(c), except that t fails the condition $t \notin \mathbf{W}_{\text{li}}$. However, that latter condition is not used in the proof that if $t \neq t'$ and t' satisfy the other conditions of Case 3(c), then $s' \neq s$, where s' is the transformation obtained from t' by the rules of s. Thus s is also different from the transformations in $f(T_n)$ from Case 3(c).

Because f is injective, $s \notin f(T_n)$, $s \in \mathbf{W}_{\text{li}}$ and $f(T_n) \subseteq \mathbf{W}_{\text{li}}$, the bound $n^{n-1} + n - 1$ cannot be reached if the length of the maximal-length chains is not 2. □

Proposition 5 *For $n \geq 4$, the minimal number of generators of the transition semigroup $\mathbf{W}_{\text{li}}$ is 5.*

Proof We need a generator, say e, that maps 0 to a state in $Q_n \setminus \{0\}$. Since all such transformations in $\mathbf{W}_{\text{li}}$ are constant transformations, e is also constant.

Let U be the set of all transformations that map $Q_n \setminus \{0\}$ to $Q_n \setminus \{0\}$ and fix 0. The transition semigroup $\mathbf{W}_{\text{li}}$ contains U. If a transformation $t \in U$ would be generated by a generator g mapping a state q from $Q_n \setminus \{0\}$ to 0, then g must be used together with some constant generator s to map 0 back to a state p in $Q_n \setminus \{0\}$. Then $0t = (0g)s = p$, since s is constant; hence t does not fix 0, which is a contradiction. Hence, all the transformations in U must be generated by generators in U.

When restricted to $Q_n \setminus \{0\}$, the set U forms the full transformation semigroup with $n - 1 \geq 3$ states. So by Remark 1, from Proposition 2 we need at least three generators for this semigroup, say a, b, and c.

Finally, T_n contains transformations mapping some states from $Q_n \setminus \{0\}$ to 0, so we need one more generator, say d, mapping a state from $Q_n \setminus \{0\}$ to 0. □

We are finally in a position to prove our main theorem of this section.

Theorem 4 (**Left Ideals, Suffix-Closed Languages**) *Suppose that $L \subseteq \Sigma^*$ and $\kappa(L) = n$. If L is a left ideal or a suffix-closed language, then $\sigma(L) \le n^{n-1} + n - 1$. This bound is tight for $n = 1$ if $|\Sigma| \ge 1$, for $n = 2$ if $|\Sigma| \ge 3$, for $n = 3$ if $|\Sigma| \ge 4$, and for $n \ge 4$ if $|\Sigma| \ge 5$. Moreover, the sizes of the alphabet cannot be reduced.*

Proof If L is a left ideal, then $\sigma(L_n) \le n^{n-1} + n - 1$ by Lemma 5. By Lemma 3 the languages of Definition 2 meet this bound. It is easy to verify that the size of the alphabet cannot be reduced if $n \le 3$. For $n \ge 4$, by Theorem 3 only languages L whose quotient automata have transition semigroups isomorphic to $\mathbf{W}_{\text{li}}$ meet the bound, and by Proposition 5 $\mathbf{W}_{\text{li}}$ requires 5 generators. □

6 Two-Sided Ideals

If a language L is a right ideal, then $L = L\Sigma^*$ and L has exactly one final quotient, namely Σ^*; hence this also holds for two-sided ideals. For $n \ge 3$, in a two-sided ideal every maximal chain is of length at least 3: it starts with L, every quotient contains L and is contained in Σ^*.

6.1 Lower Bound

We now show that the syntactic complexity of the following DFA of a two-sided ideal is $n^{n-2} + (n-2)2^{n-2} + 1$.

Definition 3 (**Witness: Two-Sided Ideals**) For $n \ge 4$, define the DFA $\mathcal{W}_n = (Q_n, \Sigma_{\mathcal{W}}, \delta_{\mathcal{W}}, 0, \{n-1\})$, where $\Sigma_{\mathcal{W}} = \{a, b, c, d, e, f\}$, $a\colon (1, \dots, n-2)$, $b\colon (1, 2)$, $c\colon (n-2 \to 1)$, $d\colon (n-2 \to 0)$, $e\colon Q_{n-1} \to 1$, and $f\colon (1 \to n-1)$. For $n = 4$, inputs a and b coincide, and we can use $\Sigma_{\mathcal{W}} = \{a, c, d, e, f\}$. Also, let $\mathcal{W}_3 = (Q_3, \{a, b, c\}, \delta_{\mathcal{W}}, 0, \{2\})$, where $a\colon (1 \to 2)(0 \to 1)$, $b\colon (1 \to 0)$, and $c\colon \mathbf{1}$, and let $\mathcal{W}_2 = (Q_2, \{a, b\}, \delta_{\mathcal{W}}, 0, \{1\})$, where $a\colon (0 \to 1)$, and $b\colon \mathbf{1}$. Finally, let $L_n = L(\mathcal{W}_n)$.

The structure of the DFA of Definition 3 is shown in Fig. 6 for $n \ge 4$.

Lemma 6 *For $n \ge 2$, the DFA of Definition 3 is minimal, accepts a two-sided ideal, and its transition semigroup has size $n^{n-2} + (n-2)2^{n-2} + 1$. In particular, in contains all transformations of Q_n that*

1. *fix 0 and $n-1$,*
2. *map $S \cup \{n-1\}$ to $n-1$ and $Q_n \setminus (\{S\} \cup \{n-1\})$ to i, for all $S \subseteq \{1, \dots, n-2\}$ and $i \in \{1, \dots, n-2\}$,*
3. *map Q_n to $n-1$.*

Proof For $n = 2$, the DFA $\mathcal{W}_2$ has only two states 0 and 1, and is obviously minimal. Also, $L(\mathcal{W}_2) = \{a, b\}^*a\{a, b\}^*$ is a two-sided ideal. The set S is empty, and W_2 contains all transformations of types 1 and 3. Finally, $\mathcal{W}_2$ meets the bound 2.

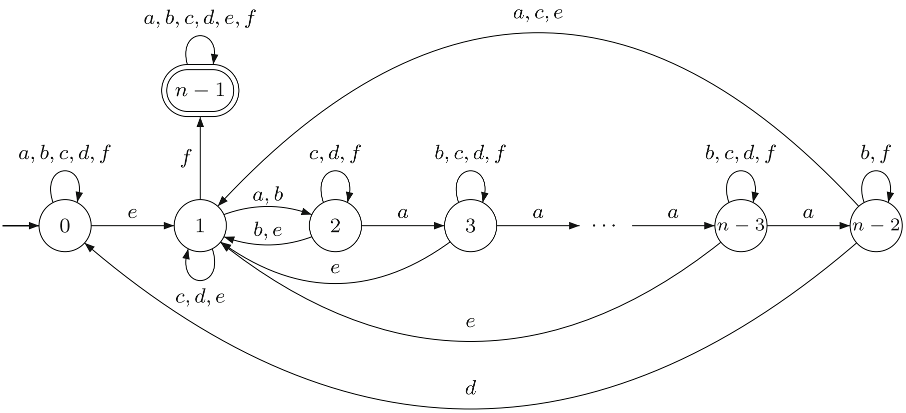

Fig. 6 Quotient DFA of a two-sided ideal with $n^{n-2} + (n-2)2^{n-2} + 1$ transformations

For $i = 1, \ldots, n-2$, state i is the only non-final state that accepts $a^{n-1-i} f$; hence all these states are distinguishable. State 0 is distinguishable from these states, because it does not accept any words in $a^* f$. Hence $\mathcal{W}_n$ is minimal. The proof that $\mathcal{W}_n$ is a left ideal is like that in Lemma 3. Since $n-1$ is the only final state and it accepts Σ^* L_n is a right ideal. Hence it is two-sided.

For $n = 3$, $\mathcal{W}_3$ meets the bound 6 with the transition semigroup consisting of the transformations [0, 0, 2], [0, 1, 2], [0, 2, 2], [1, 1, 2], [1, 2, 2], and [2, 2, 2].

From now on we may assume that $n \geq 4$. In $\mathcal{W}_n$, the transformations induced by a, b, and c restricted to $Q_n \setminus \{0, n-1\}$ generate all the transformations of the states $1, \ldots, n-2$. When restricted to $Q_n \setminus \{n-1\}$, together with the transformation of d, they generate all $(n-1)^{n-2}$ transformations that fix 0: Let t be such a transformation mapping a subset $S \subseteq Q_n \setminus \{n-1\}$ to 0. First, using a, b, c, we can map S to $n-2$ and $n-2$ to a state from S (unless $n-2 \in S$). Then we apply d. Finally, using a, b, c, we can map $n-2$ to the original state, and the remaining states as in t.

In the same way, together with the transformation f, we have all n^{n-2} transformations of Q_n that fix 0 and $n-1$.

For any subset $S \subseteq \{1, \ldots, n-2\}$, there is a transformation – induced by a word w_S, say – that maps S to $n-1$ and fixes $Q_n \setminus S$. Then the words of the form $w_S e a^i$, for $i \in \{0, \ldots, n-3\}$, induce all transformations that map $S \cup \{n-1\}$ to $n-1$ and $Q_n \setminus (S \cup \{n-1\})$ to $i+1$. There are 2^{n-2} subsets S, and there are $n-2$ possibilities for i. Hence there are $(n-2)2^{n-2}$ transformations of this type. There is also the constant transformation $ef\colon (Q_n \to n-1)$, which yields the total number claimed. □

6.2 Upper Bound

We consider a minimal DFA $\mathcal{D}_n = (Q_n, \Sigma_{\mathcal{D}}, \delta_{\mathcal{D}}, 0, \{n-1\})$ of an arbitrary two-sided ideal with n quotients, and let T_n be the transition semigroup of $\mathcal{D}_n$. We also deal with the witness DFA $\mathcal{W}_n = (Q_n, \Sigma_{\mathcal{W}}, \delta_{\mathcal{W}}, 0, \{n-1\})$ of Definition 3 with transition semigroup $\mathbf{W}_{2\mathrm{i}}$.

Lemma 7 *If $n \geq 4$ and a maximal-length chain in $\mathcal{D}_n$ strictly ordered by $\prec$ has length 3, then $|T_n| \leq n^{n-2} + (n-2)2^{n-2} + 1$, and T_n is a subsemigroup of* $\mathbf{W}_{2\mathrm{i}}$.

Proof Consider an arbitrary transformation $t \in T_n$; then $(n-1)t = n-1$. If $0t = 0$, then any state not in $\{0, n-1\}$ can possibly be mapped by t to any one of the n states; hence there are at most n^{n-2} such transformations.

If $0t \neq 0$, then $0 \prec 0t$. Consider any state $q \notin \{0, 0t\}$; since $\mathcal{D}_n$ is minimal, q must be reachable from 0 by some transformation s, that is, $q = 0s$. If $0st \notin \{0t, n-1\}$, then $0t \prec 0st$ by Remark 4. But then we have the chain $0 \prec 0t \prec 0st \prec n-1$ of length 4, contradicting our assumption. Hence we must have either $0st = 0t$, or $0st = n-1$. For a fixed $0t$, a subset of the states in $Q_n \setminus \{0, n-1\}$ can be mapped to $0t$ and the remaining states in $Q_n \setminus \{0, n-1\}$ to $n-1$, thus giving 2^{n-2} transformations. Since there are $n-2$ possibilities for $0t$, we obtain the second part of the bound. Finally, all states can be mapped to $n-1$.

By Lemma 6 all of the above-mentioned transformations are in $\mathbf{W}_{2\mathrm{i}}$.

□

Lemma 8 (**Two-Sided Ideals, Factor-Closed Languages**) *If L is a two-sided ideal or a factor-closed language with $n \geq 2$ quotients, then its syntactic complexity is less than or equal to $n^{n-2} + (n-2)2^{n-2} + 1$.*

Proof It suffices to prove the result for two-sided ideals, since factor-closed languages are their complements.

As we did for left ideals, we show that $|T_n| \leq |\mathbf{W}_{2\mathrm{i}}|$, by constructing an injective function $f \colon T_n \to \mathbf{W}_{2\mathrm{i}}$.

We have $q \preceq n-1$ for all $q \in Q_n$, and $n-1$ is a fixed point of every transformation in T_n and $\mathbf{W}_{2\mathrm{i}}$.

For a transformation $t \in T_n$, consider the cases shown in Fig. 7.

We now prove the lemma for each of these cases.

Case 1: $t \in \mathbf{W}_{2\mathrm{i}}$.

The proof is the same as that of Case 1 of Lemma 5.

Case 2: $t \notin \mathbf{W}_{2\mathrm{i}}$, and $0t^2 \neq 0t$.

Let $0t = p \prec \cdots \prec pt^k = pt^{k+1}$ be the chain defined from p.

- **(a):** $pt^k \neq n-1$.

The proof is the same as that of Case 2 of Lemma 5.

- **(b):** $pt^k = n-1$ and $k \geq 2$.

Let $f(t) = s$, where s is the transformation shown in Fig. 7, Case 2(b), and defined by

$$0s = 0, \quad pt^i s = pt^{i-1} \text{ for } 1 \leq i \leq k-1, \quad ps = n-1,$$
$$qs = qt \text{ for the other states } q \in Q_n.$$

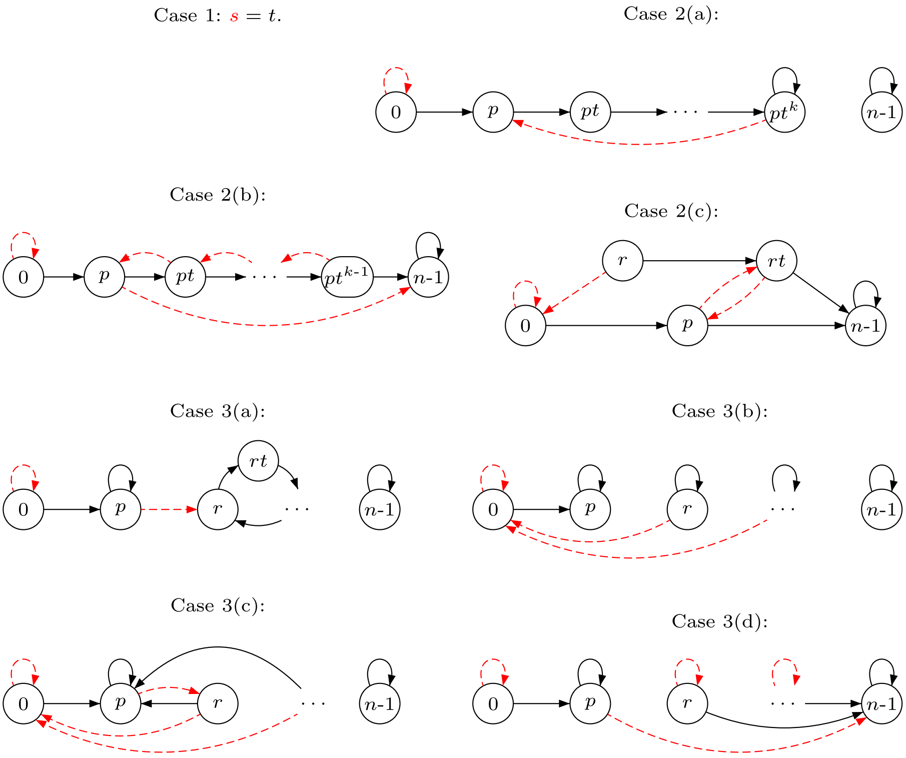

Fig. 7 Map of the cases in the proof of Lemma 8. The transitions of t are represented by *solid lines*, and the modified transitions of s by *dashed red lines*

By Lemma 6, $s \in \mathbf{W}_{2\mathrm{i}}$. We have $pt \succ p$, $pts = p$, and $ps = n-1$. By Proposition 4, $pts \succeq ps$, that is, $p \succeq n-1$, which contradicts the fact that $k \geq 2$ (so $0t = p \neq n-1$), and $q \preceq n-1$ for all $q \in Q_n$. Thus s is not in T_n, and so it is different from the transformations of Case 1.

Observe that s does not have a cycle with states strictly ordered by $\prec$, since no state from $\{0, p, pt, \ldots, pt^{k-1}\}$ can be in a cycle, and t cannot have such a cycle with ordered states by Proposition 4. Hence s is different from the transformations of Case 2(a).

In s, there is a unique state q such that $qs = n-1$ and for which there exists a state r such that $r \succ q$ and $rs = q$, and that this state q must be p. Indeed, if $q \neq p$, then $qt = qs = n-1$ by the definition of s. From $r \succ q$, we have $rt \succeq qt = n-1$; hence $rs = rt = n-1$ and $rs \neq q$ – a contradiction. Hence $q = p$.

By a similar argument, we show that there exists a unique state q such that $q \succ p$, and $qs = p$, and that this state q must be pt. If $q \neq pt$ then $qs = qt$. But $q \succ qt$ and $p = qt \succeq qt^2 = pt$ contradicts that $p \prec pt$. Continuing in this way for $pt^2, \ldots, pt^{k-1}$ we show that there is a unique chain $pt^{k-1} \xrightarrow{s} \ldots \xrightarrow{s} pt \xrightarrow{s} p$.

If $t' \neq t$ is another transformation satisfying the conditions of this case, we define s' like s. Now suppose that $s = f(t) = f(t') = s'$. Since we have a unique state p

such that $ps = n-1$ for which there exists a state r such that $r \succ p$ and $rs = p$, we have $0t = 0t' = p$. Also the chain of states $p, pt, pt^2, \dots, pt^{k-1}$ is unique in s and s' as we have shown above; so $pt^i = pt'^i$ for $i = 1, \dots, k-1$. Since the other states are mapped by s exactly as by t and t', we have $t = t'$.

- **(c):** $pt = n-1$.

Let $P = \{0, p, n-1\}$. We have $n \geq 4$, as otherwise $t \in \mathbf{W}_{2\mathrm{i}}$ since it is a transformation of type 2 from Lemma 6. So there must be a state $r \notin P$; let r be chosen arbitrarily. If $p \prec r$ for all $r \notin P$, then $n-1 = pt \preceq rt$; hence $rt = n-1$ for all such r, and $qt \in \{p, n-1\}$ for all $q \in Q_n$. By Lemma 6, there is a transformation in $\mathbf{W}_{2\mathrm{i}}$ that maps $S \cup \{n-1\}$ to $n-1$, and $Q_n \setminus (S \cup \{n-1\})$ to p for any $S \subseteq \{1, \dots, n-2\}$. Thus $t \in \mathbf{W}_{2\mathrm{i}}$ – a contradiction.

In view of the above, there must exist a state $r \notin P$ such that $p \not\preceq r$. By Remark 4, we have $p \preceq rt$ and of course $rt \preceq n-1$. If rt is p or $n-1$ for all $r \notin P$, we again have the situation described above, showing that $t \in \mathbf{W}_{2\mathrm{i}}$. Hence there must exist an $r \notin P$ such that $p \not\preceq r$ and $p \prec rt \prec n-1$.

Also we claim that t does not have a cycle. Indeed, if $p \preceq q$, then q is mapped to $n-1$; if $p \not\preceq q$, then q is mapped to a state $qt \succeq p$ and again q cannot be in a cycle since the chain starting with q ends in $n-1$.

Let $f(t) = s$, where s is the transformation shown in Fig. 7, Case 2(c), and defined by

$$0s = 0, \quad ps = rt, \quad (rt)s = p, \quad rs = 0,$$
$$qs = qt \text{ for the other states } q \in Q_n.$$

Since s fixes both 0 and $n-1$, it is in $\mathbf{W}_{2\mathrm{i}}$ by Lemma 6. But s is not in T_n, as we have the cycle (p, rt) with $p \prec rt$, which would contradict Proposition 4. So s is different from the transformations of Case 1. Since s maps a state other than 0 to 0, it is different from the transformations of Cases 2(a) and 2(b).

Observe that t does not map any state to 0; otherwise, if $qt = 0$ for some q, then $0 \prec p$ implies $q \prec 0$ by Proposition 4, which contradicts that $0 \prec q$ from Remark 4. Consequently, in s there is the unique state $r \neq 0$ mapped to 0. Also, as t does not contain a cycle, the only cycle in s must be (p, rt).

If $t' \neq t$ is another transformation satisfying the conditions of this case, we define s' like s. Now suppose that $s = f(t) = f(t') = s'$. Because both s and s' have the unique non-fixed point r mapped to 0, $r = r'$. Also s and s' contain the unique cycle (p, rt), $p \prec rt$. Thus $p = p'$, $pt = pt' = n-1$ and $rt = rt'$. It follows that $0t = 0t' = p$. Because $p \prec rt = rt'$, we have $(rt)t = (rt)t' = n-1$. The other states are mapped by s exactly as by t and t', and so $t = t'$.

Case 3: $t \notin \mathbf{W}_{2\mathrm{i}}$, $0t = p \neq 0$, and $pt = p$.

- **(a):** t has a cycle.

The case is analogous to that of Case 3(a) in Lemma 5.

Since t has a cycle, take a state r from the cycle; then r and rt are not comparable under $\preceq$ by Proposition 4, and $p \prec r$ by Remark 5. Let $f(t) = s$, where s is the transformation shown in Fig. 7 and defined by

$$0s = 0, \quad ps = r, \quad qs = qt \text{ for the other states } q \in Q_n.$$

The proof that s is different from the s of Case 1, 2(a), and that there is no $t' \neq t$ fitting in this case and yielding the same s, is the same as in Lemma 5.

In s there is the state p with the property that $p \prec ps$ but ps and ps^2 are not comparable under $\preceq$. Consider a transformation t' that fits in Case 2(b). Then in s' every state $q' = p't'^i$ for $0 \leq i \leq k-1$, and $q = 0$, is such that $q's'$ is comparable with $q's'^2$ under $\preceq$. So if there is such a state in s', it must be also present in $t' \in T_n$. But then $q' \prec q't'$ implies $q't' \preceq q't'^2$ by Proposition 4, so this is not possible. Thus $s \neq s'$.

For a distinction from the transformations of Case 2(c) observe that s does not map to 0 any state other than 0.

- **(b):** t has no cycles and has a fixed point $r \notin \{p, n-1\}$.

The case is analogous to that of Case 3(b) in Lemma 5.

Because $0 \prec r$ by Remark 4, $0t \preceq rt$ by Proposition 4. Since r is a fixed point of t, then $p = 0t \preceq rt = r$. Since $r \neq p$, we have $p \prec r$. Let $f(t) = s$, where s is the transformation shown in Fig. 7 and defined by

$$0s = 0, \quad qs = 0 \text{ for each fixed point } q \notin \{0, p, n-1\},$$

$$qs = qt \text{ for the other states } q \in Q_n.$$

The proof that s is different from the s of Case 1, 2(a), 3(a), and that there is no $t' \neq t$ fitting in this case and yielding the same s, is the same as in Lemma 5.

Since s maps to 0 a state other than 0, this case is distinct from Case 2(b). Because t does not have a cycle, and no state q mapped to 0 can be in a cycle in s, it follows that s does not have a cycle. Thus s is different from the transformations of Case 2(c).

- **(c):** t has neither a cycle nor a fixed point $r \notin \{p, n-1\}$, and has a state $r \succ p$ mapped to p.

The case is analogous to that of Case 3(c) in Lemma 5.

Let $f(t) = s$, where s is the transformation shown in Fig. 7 and defined by

$$0s = 0, \quad ps = r, \quad qs = 0 \text{ for each } q \succ p \text{ such that } qt = p,$$

$$qs = qt \text{ for the other states } q \in Q_n.$$

The proof that s is different from the s of Case 1, 2(a), 3(a), 3(b), and that there is no $t' \neq t$ fitting in this case and yielding the same s, is the same as in Lemma 5.

Since s maps to 0 a state other than 0, this case is distinct from Case 2(b). In s, 0 cannot be in a cycle, no state $q \succ p$ mapped to 0 can be in a cycle and p cannot be in a cycle as $ps = r$ and $rs = 0$. Since the other states are mapped as in t, s does not have a cycle. Thus s is different from the transformations of Case 2(c).

- **(d):** t has no cycles, no fixed point $r \notin \{p, n-1\}$, and no state $r \succ p$ mapped to p, and has a state r such that $p \prec r \prec n-1$ that is mapped to $n-1$.

Let $f(t) = s$, where s is the transformation shown in Fig. 7, Case 3(d), and defined by

$$0s = 0, \quad qs = q \text{ for states } q \text{ such that } qt = n-1, \quad ps = n-1$$

$$qs = qt \text{ for the other states } q \in Q_n.$$

By Lemma 6, $s \in \mathbf{W}_{2\mathrm{i}}$. However, s is not in T_n, as we have a fixed point r such that $p \prec r \prec n-1$ and $ps = n-1$. So Proposition 4 yields $n-1 = ps \preceq rs = r$ – a contradiction. Thus s is different from the transformations of Case 1.

Transformation s does not have any cycles, as t does not have one in this case and fixed points q and p cannot be in a cycle. So s is different from the transformations of Cases 2(a) and 3(a). Also, since p is the unique state mapped to $n-1$ and there is no state $r \succ p$ mapped to p, s is different from the transformations of Case 2(b). For a distinction from the transformations of Cases 2(c), 3(b) and 3(c), observe that s does not map to 0 any state other than 0.

If $t' \neq t$ is another transformation satisfying the conditions of this case, we define s' like s. Now suppose that $s = f(t) = f(t') = s'$. Observe that t does not have a fixed point other than $n-1$. So for every fixed point $q \notin \{0, n-1\}$ of s we have $qt = qt' = n-1$. Also, since p is the unique state mapped to $n-1$ in s, $0t = 0t' = p$ and $pt = pt' = p$. The other states are mapped by s as by t and t'; so $t = t'$.

- **All cases are covered:**

We need to ensure that any transformation t fits in at least one case. It is clear that t fits in Case 1 or 2 or 3. Any transformation from Case 2 fits in Case 2(a) or 2(b) or 2(c). For Case 3, it is sufficient to show that if (i) $t \notin \mathbf{W}_{2\mathrm{i}}$ does not contain a fixed point $r \notin \{p, n-1\}$, and (ii) there is no state r, $p \prec r \prec n-1$, mapped to p or $n-1$, then t has a cycle.

If there is no state r such that $p \prec r \prec n-1$, then $qt \in \{p, n-1\}$ for all $q \in Q_n$, since $qt \succeq p$. By the proof of Lemma 6 in $\mathbf{W}_{2\mathrm{i}}$ for any $S \subseteq Q_n \setminus \{n-1\}$ there are all transformations that map $S \cup \{n-1\}$ to $n-1$, and the other states $Q_n \setminus (S \cup \{n-1\})$ to any state from Q_n; thus $t \in \mathbf{W}_{2\mathrm{i}}$ – a contradiction.

So let t be a transformation that fits in Case 3 and satisfies (i) and (ii), and let r be some state such that $p \prec r \prec n-1$. Consider the sequence $r, rt, rt^2, \ldots$. By Remark 5, $p \preceq rt^i$ for all $i \geq 0$. If $rt^k \in \{p, n-1\}$ for some $k \geq 1$, let k be the smallest such number, then $rt^{k-1} \notin \{p, n-1\}$; we have $p \prec rt^{k-1} \prec n-1$ and $(rt^{k-1})t \in \{p, n-1\}$, contradicting (ii).

Since p and $n-1$ are the only fixed points by (i), we have $rt^i \neq rt^{i-1}$. Since there are finitely many states, $rt^i = rt^j$ for some i and j such that $0 \leq i < j-1$, and so the states $rt^i, rt^{i+1} \ldots, rt^j = rt^i$ form a cycle. □

Theorem 5 *If T_n has size $n^{n-2} + (n-2)2^{n-2} + 1$, then $T_n = \mathbf{W}_{2\mathrm{i}}$.*

Proof The proof is very similar to that of Theorem 3.

Consider a maximal-length chain of states strictly ordered by $\prec$ in $\mathcal{D}_n$. If its length is 3, then by Lemma 7 T_n is a subsemigroup of $\mathbf{W}_{2\mathrm{i}}$. Thus only $T_n = \mathbf{W}_{2\mathrm{i}}$ reaches the bound.

If there is a chain of length 4, then there are states p and r such that $0 \prec p \prec r \prec n-1$. Let $R = \{q \in Q_n \setminus \{n-1\} \mid p \prec q\}$. To define s we use the transformation $t = (Q_n \setminus \{n-1\} \to p)$ as an auxiliary transformation. Note that t fits in Case 3(c) in the proof of the proof of Lemma 8 except that $t \in \mathbf{W}_{2\mathrm{i}}$. We define s from t according to the rules of Case 3(c):

$$0s = 0, \quad ps = r, \quad qs = 0 \text{ for each} q \in R,$$
$$qs = qt = p \text{ for the other states} q.$$

By Lemma 6 (transformations of type 1), $s \in \mathbf{W}_{2\mathrm{i}}$.

Let f be the injective function from the proof of Lemma 8. It remains to be shown that there is no transformation $t' \in T_n$ such that $s = f(t')$. The proof that s is different from the transformations $f(t')$ of Cases 1, 2(a), 2(b), 2(c), 3(a), and 3(b) is exactly the same as the corresponding proof in Case 3(c) following the definition of s. The proof that s is different from the transformations $t' \in T_n$ in Case 3(c) is exactly the same as the corresponding proof in Theorem 3. It remains to show that that there is no $t' \in T_n$ in Case 3(d) such that $s = f(t')$. Indeed, $f(t')$ from Case 3(d) does not map any state to 0 other than 0, while we have $rs = 0$. So s is also different from these transformations.

Because f is injective, $s \notin f(T_n)$, $s \in \mathbf{W}_{2\mathrm{i}}$ and $f(T_n) \subseteq \mathbf{W}_{2\mathrm{i}}$, the bound $n^{n-1} + n - 1$ cannot be reached if the length of the maximal-length chains is not 3. □

Proposition 6 *For $n \geq 4$, the minimal number of generators of the transition semigroup $\mathbf{W}_{2\mathrm{i}}$ is 6.*

Proof Transition semigroup $\mathbf{W}_{2\mathrm{i}}$ contains all transformations of Q_{n-1} to Q_{n-1} that fix $n-1$. Since every transformation in $\mathbf{W}_{2\mathrm{i}}$ fixes $n-1$, they must be also generated only by transformations of this form, as otherwise a generated transformation would map a state from Q_{n-1} to $n-1$, as $n-1$ is always fixed. When restricted to Q_{n-1}, these transformations form the largest transformation semigroup of a left ideal $\mathbf{W}_{1\mathrm{i}}$ with $n-1 \geq 4$ states. So by Remark1, from Proposition 5 we know that they require 5 generators. These generators do not map any state from $Q_n \setminus \{n-1\}$ to $n-1$, hence, we need one more generator which maps a state from Q_{n-1} to $n-1$. □

We are now in a position to prove our main theorem of this section.

Theorem 6 (**Two-Sided Ideals, Factor-Closed Languages**) *Suppose that $L \subseteq \Sigma^*$ and $\kappa(L) = n > 1$. If L is a two-sided ideal or a factor-closed language, then $\sigma(L) \leq n^{n-2} + (n-2)2^{n-2} + 1$. This bound is tight for $n = 2$ if $|\Sigma| \geq 2$, for $n = 3$ if $|\Sigma| \geq 3$, for $n \geq 4$ if $|\Sigma| \geq 5$, and for $n \geq 5$ if $|\Sigma| \geq 6$. Moreover, the sizes of the alphabet cannot be reduced.*

Proof This follows from Lemmas 6 and 8. It is easy to verify that the size of the alphabet cannot be reduced if $n \leq 4$. For $n \geq 5$, by Theorem 5 only languages L whose quotient automaton has transition semigroup isomorphic to $\mathbf{W}_{2\mathrm{i}}$ meet the bound, and by Proposition 6, transition semigroup $\mathbf{W}_{2\mathrm{i}}$ requires 6 generators. □

7 Conclusions

We have found tight upper bounds on the syntactic complexity of right, left, and two-sided ideals. We have shown that in each of the three cases the maximal transition semigroup is unique.

In our proof for left and two-sided ideals we exhibited an injective function from the transition semigroup of a minimal DFA of an arbitrary left, right, two-sided ideal language to the transition semigroup of the witness DFA attaining the upper bound for these languages. This approach is generally applicable for other subclasses of regular languages. For example, in [12] we have used this method to establish the upper bound for suffix-free languages.

Acknowledgements This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant No. OGP000087, the National Science Centre, Poland under project number 2013/09/N/ST6/01194, an NSERC Postgraduate Scholarship, and a Graduate Award from the Department of Computer Science, University of Toronto. It was completed during the internship of Marek Szykuła at the University of Waterloo, which was co-financed by the European Union under the European Social Fund's project "International computer science and applied mathematics for business study programme at the University of Wrocław".

References

1. Ang, T., Brzozowski, J.A.: Languages convex with respect to binary relations, and their closure properties. Acta Cybernet. **19**(2), 445–464 (2009)
2. Brzozowski, J.A.: Quotient complexity of regular languages. J. Autom. Lang. Comb. **15**(1/2), 71–89 (2010)
3. Brzozowski, J.A.: In search of the most complex regular languages. Int. J. Found. Comput. Sc. **24**(6), 691–708 (2013)
4. Brzozowski, J.A., Jirásková, G., Li, B.: Quotient complexity of ideal languages. Theoret. Comput. Sci. **470**, 36–52 (2013)
5. Brzozowski, J.A., Jirásková, G., Zou, C.: Quotient complexity of closed languages. Theory Comput. Syst. **54**, 277–292 (2014)
6. Brzozowski, J.A., Li, B.: Syntactic complexity of R- and J-trivial languages. Internat. J. Found. Comput. Sci. **16**(3), 547–563 (2005)
7. Brzozowski, J.A., Li, B., Liu, D.: Syntactic complexities of six classes of star-free languages. J. Autom. Lang. Comb. **17**, 83–105 (2012)
8. Brzozowski, J.A., Li, B., Ye, Y.: Syntactic complexity of prefix-, suffix-, bifix-, and factor-free regular languages. Theoret. Comput. Sci. **449**, 37–53 (2012)
9. Brzozowski, J.A., Shallit, J., Xu, Z.: Decision problems for convex languages. Inf. Comput. **209**, 353–367 (2011)
10. Brzozowski, J.A., Szykuła, M.: Large aperiodic semigroups. Int. J. Found. Comput. Sc. **26**(7), 913–931 (2015)
11. Brzozowski, J.A., Szykuła, M.: Upper bounds on syntactic complexity of left and two-sided ideals. In: Shur, A.M., Volkov, M.V. (eds.) DLT 2014, LNCS, vol. 8633, pp. 13–24. Springer (2014)

12. Brzozowski, J.A., Szykuła, M.: Upper bound for syntactic complexity of suffix-free languages. In: Okhotin, A., Shallit, J. (eds.) DCFS 2015, LNCS, vol. 9118, pp. 33–45. Springer (2015). Full paper to appear in Information and Computation
13. Brzozowski, J.A., Tamm, H.: Theory of átomata. Theoret. Comput. Sci. **539**, 13–27 (2014)
14. Brzozowski, J.A., Ye, Y.: Syntactic complexity of ideal and closed languages. In: Mauri, G., Leporati, A. (eds.) DLT 2011, LNCS, vol. 6795, pp. 117–128. Springer (2011)
15. Câmpeanu, C., Culik, K. II., Salomaa, K., Yu, S.: State complexity of basic operations on finite languages. In: Boldt, O., Jürgensen, H. (eds.) Automata implementation, LNCS, vol. 2214, pp. 60–70. Springer (2001)
16. Crochemore, M., Hancart, C.: Automata for pattern matching. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages, vol. 2, pp. 399–462. Springer (1997)
17. Han, Y.S., Salomaa, K.: State complexity of basic operations on suffix-free regular languages. Theoret. Comput. Sci. **410**(27–29), 2537–2548 (2009). doi:10.1016/j.tcs.2008.12.054
18. Han, Y.S., Salomaa, K., Wood, D., Ésik, Z., Fülöp, Z.: Operational state complexity of prefix-free regular languages. In: Automata, Formal Languages, and Related Topics, pp. 99–115. University of Szeged, Hungary (2009)
19. Holzer, M., König, B.: On deterministic finite automata and syntactic monoid size. Theoret. Comput. Sci. **327**, 319–347 (2004)
20. Krawetz, B., Lawrence, J., Shallit, J.: State complexity and the monoid of transformations of a finite set. Internat. J. Found. Comput. Sci. **16**(3), 547–563 (2005)
21. de Luca, A., Varricchio, S.: Some combinatorial properties of factorial languages. In: Capocelli, R. (ed.) Sequences: Combinatorics, compression, security, and transmission, pp. 258–266. Springer (1990)
22. Maslov, A.N.: Estimates of the number of states of finite automata. Dokl. Akad. Nauk SSSR **194** 1266–1268, (Russian) (1970). English translation: Soviet Math. Dokl. **11** (1970), 1373–1375
23. McNaughton, R., Papert, S.A.: Counter-free Automata (M.I.T. Research Monograph No. 65). The MIT Press (1971)
24. Myhill, J.: Finite automata and representation of events. Wright Air Development Center Technical Report. **57–624** (1957)
25. Nerode, A.: Linear automaton transformations. Proc. Amer. Math. Soc. **9**, 541–544 (1958)
26. Paz, A., Peleg, B.: Ultimate-definite and symmetric-definite events and automata. J. ACM **12**(3), 399–410 (1965). doi:10.1145/321281.321292
27. Perrin, D.: Finite automata. In: Van Leewen, J. (ed.) Handbook of Theoretical Computer Science, vol. B, pp. 1–57. Elsevier (1990)
28. Piccard, S.: Sur les bases du group symétrique et du groupe alternant. Commentarii Mathematici Helvetici **11**(1), 1–8 (1938)
29. Pin, J.E.: Syntactic semigroups. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages, Vol. 1: Word, Language, Grammar, pp. 679–746. Springer (1997)
30. Sierpiński, W.: Sur les suites infinies de fonctions définies dans les ensembles quelconques. Fund. Math. **24**, 209–212 (1935)
31. Szykuła, M., Wittnebel, J.: Syntactic complexity of bifix-free languages. In: Carayol, A., Nicaud, C. (eds.) CIAA 2017, LNCS, vol. 10329, pp. 201–212. Springer (2017). Full paper at arXiv:1604.06936
32. Yu, S.: Regular languages. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages, pp. 41–110. Springer (1997)
33. Yu, S.: State complexity of regular languages. J. Autom. Lang. Comb. **6**, 221–234 (2001)