# Syntactic Complexity of Prefix-, Suffix-, and Bifix-Free Regular Languages*

Janusz Brzozowski[1], Baiyu Li[1], and Yuli Ye[2]

[1] David R. Cheriton School of Computer Science, University of Waterloo
Waterloo, ON, Canada N2L 3G1
{brzozo,b5li}@uwaterloo.ca
[2] Department of Computer Science, University of Toronto
Toronto, ON, Canada M5S 3G4
y3ye@cs.toronto.edu

**Abstract.** The syntactic complexity of a regular language is the cardinality of its syntactic semigroup. The syntactic complexity of a subclass of the class of regular languages is the maximal syntactic complexity of languages in that class, taken as a function of the state complexity $n$ of these languages. We study the syntactic complexity of prefix-, suffix-, and bifix-free regular languages. We prove that $n^{n-2}$ is a tight upper bound for prefix-free regular languages. We present properties of the syntactic semigroups of suffix- and bifix-free regular languages, and conjecture tight upper bounds on their size.

**Keywords:** bifix-free, finite automaton, monoid, prefix-free, regular language, semigroup, suffix-free, syntactic complexity.

## 1 Introduction

A language is *prefix-free* (respectively, *suffix-free*) if it does not contain any pair of words such that one is a proper prefix (respectively, suffix) of the other. It is *bifix-free* if it is both prefix- and suffix-free. We refer to prefix-, suffix-, and bifix-free languages as *free* languages. Nontrivial prefix-, suffix-, and bifix-free languages are also known as prefix, suffix, and bifix codes, respectively [1] and, as such, have many applications in areas such as cryptography, data compression, and information processing.

The *state complexity* of a regular language is the number of states in the minimal deterministic finite automaton (DFA) recognizing that language. An equivalent notion is that of *quotient complexity,* which is the number of left quotients of the language. State complexity of regular operations has been studied quite extensively: for surveys of this topic and lists of references we refer the reader to [2,19]. With regard to free regular languages, Han, Salomaa and Wood [8]

* This work was supported by the Natural Sciences and Engineering Research Council of Canada under grant No. OGP0000871 and a Postgraduate Scholarship, and by a Graduate Award from the Department of Computer Science, University of Toronto.

M. Holzer, M. Kutrib, and G. Pighizzini (Eds.): DCFS 2011, LNCS 6808, pp. 93–106, 2011.

examined prefix-free regular languages, and Han and Salomaa [7] studied suffix-free regular languages. Bifix-, factor-, and subword-free regular languages were studied by Brzozowski, Jirásková, Li, and Smith [3].

The notion of quotient complexity can be derived from the Nerode congruence [14], while the Myhill congruence [13] leads to the syntactic semigroup of a language and to its *syntactic complexity*, which is the cardinality of the syntactic semigroup. It was pointed out in [4] that syntactic complexity can be very different for regular languages with the same quotient complexity. Thus, for a fixed $n$, languages with quotient complexity $n$ may possibly be distinguished by their syntactic complexities.

In contrast to state complexity, syntactic complexity has not received much attention. In 1970 Maslov [12] dealt with the problem of generators of the semigroup of all transformations in the setting of finite automata. In 2003–2004, Holzer and König [9], and independently, Krawetz, Lawrence and Shallit [11] studied the syntactic complexity of automata with unary and binary alphabets. In 2010 Brzozowski and Ye [4] examined the syntactic complexity of ideal and closed regular languages. Here, we deal with the syntactic complexity of prefix-, suffix-, and bifix-free regular languages, and their complements.

Basic definitions and facts are stated in Sections 2 and 3. In Section 4 we obtain a tight upper bound for the syntactic complexity of prefix-free regular languages. In Sections 5 and 6 we study the syntactic complexity of suffix- and bifix-free regular languages, respectively, and we state conjectures about tight upper bounds for these classes. Section 7 concludes the paper.

## 2 Transformations

A *transformation* of a set $Q$ is a mapping of $Q$ into itself. In this paper we consider only transformations of finite sets, and we assume without loss of generality that $Q = \{1, 2, \ldots, n\}$. Let $t$ be a transformation of $Q$. If $i \in Q$, then $it$ is the image of $i$ under $t$. If $X$ is a subset of $Q$, then $Xt = \{it \mid i \in X\}$, and the *restriction* of $t$ to $X$, denoted by $t|_X$, is a mapping from $X$ to $Xt$ such that $it|_X = it$ for all $i \in X$. The *composition* of two transformations $t_1$ and $t_2$ of $Q$ is a transformation $t_1 \circ t_2$ such that $i(t_1 \circ t_2) = (it_1)t_2$ for all $i \in Q$. We usually drop the composition operator "$\circ$" and write $t_1t_2$ for short. An arbitrary transformation can be written in the form

$$t = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ i_1 & i_2 & \cdots & i_{n-1} & i_n \end{pmatrix},$$

where $i_k = kt$, $1 \leqslant k \leqslant n$, and $i_k \in Q$. The *domain* $\operatorname{dom} t$ of $t$ is $Q$. The *image* $\operatorname{img} t$ of $t$ is the set $\operatorname{img} t = Qt$. We also use the notation $t = [i_1, i_2, \ldots, i_n]$ for the transformation $t$ above. A transformation $t$ of $Q$ is *injective* on a subset $X$ of $Q$ if, for all $i, j \in X$, $i \neq j$, we have $it \neq jt$.

A *permutation* of $Q$ is a mapping of $Q$ *onto* itself. In other words, a permutation $\pi$ of $Q$ is a transformation where $\operatorname{img} \pi = Q$.

The *identity* transformation maps each element to itself, that is, $it = i$ for $i = 1, \ldots, n$. A transformation $t$ contains a *cycle* of length $k$ if there exist pairwise

different elements $i_1, \ldots, i_k$ such that $i_1 t = i_2, i_2 t = i_3, \ldots, i_{k-1} t = i_k$, and $i_k t = i_1$. A cycle will be denoted by $(i_1, i_2, \ldots, i_k)$. For $i < j$, a *transposition* is the cycle $(i, j)$, and $(i, i)$ is the identity. A *singular* transformation, denoted by $\binom{i}{j}$, has $it = j$ and $ht = h$ for all $h \neq i$, and $\binom{i}{i}$ is the identity. A *constant* transformation, denoted by $\binom{Q}{j}$, has $it = j$ for all $i$.

The set of all transformations of a set $Q$, denoted by $\mathcal{T}_Q$, is a finite monoid. The set of all permutations of $Q$ is a group, denoted by $\mathfrak{S}_Q$ and called the *symmetric group* of degree $n$. It was shown in [10] and [16] that two generators are sufficient to generate the symmetric group of degree $n$. In 1935 Piccard [15] proved that three transformations of $Q$ are sufficient to generate the monoid $\mathcal{T}_Q$. In the same year, Eilenberg showed that fewer than three generators are not possible, as reported by Sierpiński [18]. We refer the reader to the book of Ganyushkin and Mazorchuk [5] for a detailed discussion of finite transformation semigroups. Let $Q = \{1, \ldots, n\}$. The following are well-known facts about generators of $\mathfrak{S}_Q$ and $\mathcal{T}_Q$:

**Theorem 1 (Permutations, [10,16]).** *The symmetric group $\mathfrak{S}_Q$ of size $n!$ can be generated by any cyclic permutation of $n$ elements with any transposition. In particular, $\mathfrak{S}_Q$ can be generated by $c = (1, 2, \ldots, n)$ and $t = (1, 2)$.*

**Theorem 2 (Transformations, [15]).** *The complete transformation monoid $\mathcal{T}_Q$ of size $n^n$ can be generated by any cyclic permutation of $n$ elements together with a transposition and a "returning" transformation $r = \binom{n}{1}$. In particular, $\mathcal{T}_Q$ can be generated by $c = (1, 2, \ldots, n)$, $t = (1, 2)$ and $r = \binom{n}{1}$.*

## 3 Quotient Complexity and Syntactic Complexity

If $\Sigma$ is a non-empty finite alphabet, then $\Sigma^*$ is the free monoid generated by $\Sigma$, and $\Sigma^+$ is the free semigroup generated by $\Sigma$. A *word* is any element of $\Sigma^*$, and the empty word is $\varepsilon$. The length of a word $w \in \Sigma^*$ is $|w|$. A *language* over $\Sigma$ is any subset of $\Sigma^*$. If $w = uv$ for some $u, v \in \Sigma^*$, then $u$ is a *prefix* of $w$, and $v$ is a *suffix* of $w$. A *proper* prefix (suffix) of $w$ is a prefix (suffix) of $w$ other than $w$.

The *left quotient*, or simply *quotient,* of a language $L$ by a word $w$ is the language $L_w = \{x \in \Sigma^* \mid wx \in L\}$. For any $L \subseteq \Sigma^*$, the *Nerode congruence* [14] $\sim_L$ of $L$ is defined as follows:

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

Clearly, $L_x = L_y$ if and only if $x \sim_L y$. Thus each equivalence class of this congruence corresponds to a distinct quotient of $L$.

The *Myhill congruence* [13] $\approx_L$ of $L$ is defined as follows:

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

This congruence is also known as the *syntactic congruence* of $L$. The quotient set $\Sigma^+/\approx_L$ of equivalence classes of the relation $\approx_L$, is a semigroup called

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. The *monoid complexity* $\mu(L)$ of $L$ is the cardinality of its syntactic monoid. If the equivalence class containing $\varepsilon$ is a singleton in the syntactic monoid, then $\sigma(L) = \mu(L) - 1$; otherwise, $\sigma(L) = \mu(L)$.

A *deterministic finite automaton* (DFA) is a quintuple $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$, where $Q$ is a finite, non-empty set of *states*, $\Sigma$ is a finite non-empty *alphabet*, $\delta : Q \times \Sigma \to Q$ is the *transition function*, $q_1 \in Q$ is the *initial state*, and $F \subseteq Q$ is the set of *final states*. We extend $\delta$ to $Q \times \Sigma^*$ in the usual way. The DFA $\mathcal{A}$ accepts a word $w \in \Sigma^*$ if $\delta(q_1, w) \in F$. The set of all words *accepted* by $\mathcal{A}$ is $L(\mathcal{A})$. By the *language of a state* $q$ of $\mathcal{A}$ we mean the language $L_q$ accepted by the automaton $(Q, \Sigma, \delta, q, F)$. A state is *empty* if its language is empty.

Let $L$ be a regular language. The *quotient automaton* or *quotient DFA* of $L$ is $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$, where $Q = \{L_w \mid w \in \Sigma^*\}$, $\delta(L_w, a) = L_{wa}$, $q_1 = L_\varepsilon = L$, $F = \{L_w \mid \varepsilon \in L_w\}$. To simplify notation we write $\varepsilon$ for the language $\{\varepsilon\}$. The number $\kappa(L)$ of distinct quotients of $L$ is the *quotient complexity* of $L$. The quotient DFA of $L$ is the minimal DFA accepting $L$, and so quotient complexity is the same as state complexity, but there are advantages to using quotients [2].

In terms of automata, each equivalence class $[w]_{\sim_L}$ of $\sim_L$ is the set of all words $w$ that take the automaton to the same state from the initial state. In terms of quotients, it is the set of words $w$ that can be followed by the same quotient $L_w$. In terms of automata, each equivalence class $[w]_{\approx_L}$ of $\approx_L$ is the set of all words that perform the same transformation on the set of states.

Let $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$ be a DFA. For each word $w \in \Sigma^*$, the transition function for $w$ defines a transformation $t_w$ of $Q$ by the word $w$: for all $i \in Q$, $it_w \stackrel{\text{def}}{=} \delta(i, w)$. The set $T_\mathcal{A}$ of all such transformations by non-empty words forms a subsemigroup of $\mathcal{T}_Q$, called the *transition semigroup* of $\mathcal{A}$ [17] (p. 690). Conversely, we can use a set $\{t_a \mid a \in \Sigma\}$ of transformations to define $\delta$, and so the DFA $\mathcal{A}$. When the context is clear we simply write $a = t$, where $t$ is a transformation of $Q$, to mean that the transformation performed by $a \in \Sigma$ is $t$.

Let $\mathcal{A}$ be the quotient automaton of $L$. Then $T_\mathcal{A}$ is isomorphic to the syntactic semigroup $T_L$ of $L$, and we represent elements of $T_L$ by transformations in $T_\mathcal{A}$.

We attempt to obtain tight upper bounds on the syntactic complexity $\sigma(L) = |T_L|$ of $L$ as a function of the state complexity $\kappa(L)$ of $L$. First we consider the syntactic complexity of regular languages over a unary alphabet, where the concepts, prefix-, suffix-, and bifix-free, coincide. So we may consider only unary prefix-free regular languages $L$ with quotient complexity $\kappa(L) = n$. When $n = 1$, the only prefix-free regular language is $L = \emptyset$ with $\sigma(L) = 1$. For $n \geqslant 2$, a prefix-free regular language $L$ must be a singleton, $L = \{a^{n-2}\}$. The syntactic semigroup $T_L$ of $L$ consists of $n - 1$ transformations $t_w$ by words $w = a^i$, where $1 \leqslant i \leqslant n - 1$. Thus we have

**Proposition 3 (Unary Free Regular Languages).** *If $L$ is a unary prefix-free (suffix-free, bifix-free) regular language with $\kappa(L) = n \geqslant 2$, then $\sigma(L) = n - 1$.*

The tight upper bound for regular unary languages [9] is $n$.

We assume that $|\Sigma| \geqslant 2$ in the following sections. Since the syntactic semigroup of a language is the same as that of its complement, we deal only with prefix-free, suffix-free, and bifix-free languages. All the syntactic complexity results, however, apply also to the complements of these languages.

## 4 Prefix-Free Regular Languages

Recall that a regular language $L$ is prefix-free if and only it has exactly one accepting quotient, and that quotient is $\varepsilon$ [8].

**Theorem 4 (Prefix-Free Regular Languages).** *If $L$ is regular and prefix-free with $\kappa(L) = n \geqslant 2$, then $\sigma(L) \leqslant n^{n-2}$. Moreover, this bound is tight for $n = 2$ if $|\Sigma| \geqslant 1$, for $n = 3$ if $|\Sigma| \geqslant 2$, for $n = 4$ if $|\Sigma| \geqslant 4$, and for $n \geqslant 5$ if $|\Sigma| \geqslant n + 1$.*

*Proof.* If $L$ is prefix-free, the only accepting quotient of $L$ is $\varepsilon$. Thus $L$ also has the empty quotient, since $\varepsilon_a = \emptyset$ for $a \in \Sigma$. Let $\mathcal{A} = (Q, \Sigma, \delta, 1, n-1)$ be the quotient DFA of $L$, where $Q = \{1, 2, \ldots, n\}$ and, without loss of generality, $n - 1 \in Q$ is the only accepting state, and $n \in Q$ is the empty state. For any transformation $t \in T_L$, $(n-1)t = nt = n$. Thus we have $\sigma(L) \leqslant n^{n-2}$.

The only prefix-free regular language for $n = 1$ is $L = \emptyset$ with $\sigma(L) = 1$; here the bound $n^{n-2}$ does not apply. For $n = 2$ and $\Sigma = \{a\}$, the language $L = \varepsilon$ meets the bound. For $n = 3$ and $\Sigma = \{a, b\}$, $L = b^*a$ meets the bound. For $n \geqslant 4$, let $\mathcal{A}_n = (\{1, 2, \ldots, n\}, \{a, b, c, d_1, d_2, \ldots, d_{n-2}\}, \delta, 1, \{n-1\})$, where $a = \binom{n-1}{n}(1, 2, \ldots, n-2)$, $b = \binom{n-1}{n}(1, 2)$, $c = \binom{n-1}{n}\binom{n-2}{1}$, and $d_i = \binom{n-1}{n}\binom{i}{n-1}$ for $i = 1, 2, \ldots, n-2$. DFA $\mathcal{A}_6$ is shown in Fig. 1, where $\Gamma = \{d_1, d_2, \ldots, d_{n-2}\}$. For $n = 4$, input $a$ coincides with $b$; hence only 4 inputs are needed.

Any transformation $t \in T_L$ has the form

$$t = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n \\ i_1 & i_2 & i_3 & \cdots & i_{n-2} & n & n \end{pmatrix},$$

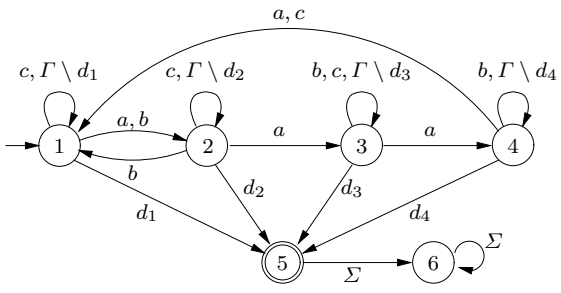


**Fig. 1.** Quotient DFA $\mathcal{A}_6$ of prefix-free regular language with 1,296 transformations

where $i_k \in \{1, 2, 3, \ldots, n\}$ for $1 \leqslant k \leqslant n-2$. There are three cases:

1. If $i_k \leqslant n-2$ for all $k$, $1 \leqslant k \leqslant n-2$, then by Theorem 2, $\mathcal{A}_n$ can do $t$.

2. If $i_k \leqslant n-1$ for all $k$, $1 \leqslant k \leqslant n-2$, and there exists some $h$ such that $i_h = n-1$, then there exists some $j$, $1 \leqslant j \leqslant n-2$ such that $i_k \neq j$ for all $k$, $1 \leqslant k \leqslant n-2$. For all $1 \leqslant k \leqslant n-2$, define $i'_k$ as follows: $i'_k = j$ if $i_k = n-1$, and $i'_k = i_k$ if $i_k \neq n-1$. Let

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n-2 & n-1 & n \\ i'_1 & i'_2 & i'_3 & i'_4 & \cdots & i'_{n-2} & n & n \end{pmatrix}.$$

By Case 1 above, $\mathcal{A}_n$ can do $s$. Since $t = sd_j$, $\mathcal{A}_n$ can do $t$ as well.

3. Otherwise, there exists some $h$ such that $i_h = n$. Then there exists some $j$, $1 \leqslant j \leqslant n-2$, such that $i_k \neq j$ for all $k$, $1 \leqslant k \leqslant n-2$. For all $1 \leqslant k \leqslant n-2$, define $i'_k$ as follows: $i'_k = n-1$ if $i_k = n$, $i'_k = j$ if $i_k = n-1$, and $i'_k = i_k$ otherwise. Let $s$ be as above but with new $i'_k$. By Case 2 above, $\mathcal{A}_n$ can do $s$. Since $t = sd_j$, $\mathcal{A}_n$ can do $t$ as well.

Therefore, the syntactic complexity of $\mathcal{A}_n$ meets the desired bound. □

We conjecture that the alphabet sizes cannot be reduced. As shown in Table 1, on p. 104, we have verified this conjecture by enumerating all prefix-free regular languages for $n \leqslant 5$ using *GAP* [6].

It was shown in [4] that for certain right, left, and two-sided ideals with maximal syntactic complexity, the reverse languages have maximal state complexity. This is also true for prefix-free languages.

**Theorem 5.** *The reverse of the prefix-free regular language accepted by automaton $\mathcal{A}_n$ of Theorem 4 restricted to $\{a, c, d_{n-2}\}$ has $2^{n-2}+1$ quotients, which is the maximum possible for a prefix-free regular language.*

*Proof.* Let $\mathcal{B}_n$ be the automaton $\mathcal{A}_n$ restricted to $\{a, c, d_{n-2}\}$. Since $L(\mathcal{A}_n)$ is prefix-free, so is $L_n = L(\mathcal{B}_n)$. We show that $\kappa(L_n^R) = 2^{n-2}+1$.

Let $\mathcal{N}_n$ be the NFA obtained by reversing $\mathcal{B}_n$. (See Fig. 2 for $\mathcal{N}_6$.) Apply the subset construction to $\mathcal{N}_n$. We prove that the following $2^{n-2}+1$ sets of states of $\mathcal{N}_n$ are reachable and distinct: $\{n-1\} \cup \{S \mid S \subseteq \{1, \ldots, n-2\}\}$.

The singleton set $\{n-1\}$ of initial states of $\mathcal{N}_n$ is reached by $\varepsilon$. From $\{n-1\}$ we reach the empty set by $a$. The set $\{n-2\}$ is reached by $d_{n-2}$ from $\{n-1\}$, and from here, $\{1\}$ is reached by $a^{n-3}$. From any set $\{1, 2, \ldots, i\}$, where $1 \leqslant i < n-2$, we reach $\{1, 2, \ldots, i, i+1\}$ by $ca^{n-3}$. Thus we reach $\{1, 2, \ldots, n-2\}$ from $\{1\}$ by

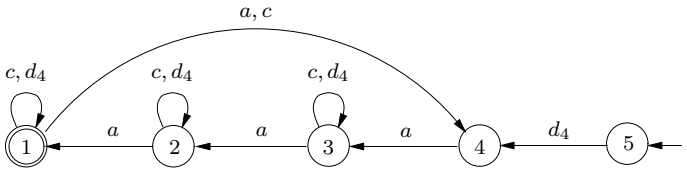


**Fig. 2.** NFA $\mathcal{N}_6$ of $L_6^R$ with quotient complexity $\kappa(L_6^R) = 17$; empty state omitted

$(ca^{n-3})^{n-3}$. Now assume that any set $S$ of cardinality $k \leqslant n-2$ can be reached; then we can get a set of cardinality $k-1$ by deleting an element $j$ from $S$ by applying $a^j d_{n-2} a^{n-2-j}$. Hence all subsets of $\{1, 2, \ldots, n-2\}$ can be reached.

The empty state accepts nothing, and the initial state $n-1$ is the only state accepting $d_{n-2}a^{n-3}$. For $1 \leqslant i \leqslant n-2$, the word $a^{i-1}$ is accepted only by state $i$ of $\mathcal{N}_n$. Suppose $S_1, S_2$, $S_1 \neq S_2$, are two non-empty subsets of $\{1, \ldots, n-2\}$. Then there exists $i \in S_1 \setminus S_2 \cup S_2 \setminus S_1$, and $w = a^i$ is accepted by one subset but not the other.

It was shown in [8] that the quotient complexity of reversal of prefix-free languages is at most $2^{n-2}+1$. Since all $2^{n-2}+1$ sets of states of $\mathcal{N}_n$ are reachable and pairwise distinguishable, $L_n$ is a witness. □

## 5 Suffix-Free Regular Languages

For any regular language $L$, a quotient $L_w$ is *uniquely reachable* if $L_w = L_x$ implies that $w = x$. It is known from [7] that, if $L$ is a suffix-free regular language, then $L = L_\varepsilon$ is uniquely reachable by $\varepsilon$, and $L$ has the empty quotient. For $n \geqslant 1$, let $Q = \{1, 2, \ldots, n\}$. Without loss of generality, for any quotient automaton of a suffix-free regular language $L$ with $Q$ as its set of states, we assume that 1 is the initial state, and $n$ is the empty state. Let

$$G_n = \{t \in \mathcal{T}_Q \mid 1 \notin \operatorname{img} t, nt = n, \text{ and } 1t = n \text{ or } 1t \neq it \text{ for } i \neq 1\}.$$

Note that $G_n$ is not a semigroup for $n \geqslant 3$ because $s = [2, 3, 3, \ldots, 3, n] \in G_n$ but $s^2 = [3, 3, 3, \ldots, 3, n] \notin G_n$. Let $g(n) = |G_n|$.

**Proposition 6.** *If $L$ is a regular language with quotient complexity $n$ and syntactic semigroup $T_L$, then the following hold:*

1. *If $L$ is suffix-free, then $T_L$ is a subset of $G_n$.*
2. *If $L$ is suffix-free and $n \geqslant 2$, then $\sigma(L) \leqslant g(n) = (n-1)^{n-2} + (n-2)^{n-1}$.*
3. *If $L$ has only one accepting quotient, and $T_L \subseteq G_n$, then $L$ is suffix-free.*

*Proof.* 1. Let $L$ be suffix-free, and let $\mathcal{A}_n = (Q, \Sigma, \delta, 1, F)$ be its quotient DFA. Consider an arbitrary $t \in T_L$. Since the quotient $L$ is uniquely reachable, $it \neq 1$ for all $i \in Q$. Since the quotient $L_n$ is empty, $nt = n$. Furthermore, since $L$ is suffix-free, for any two quotients $L_v$ and $L_{uv}$, where $u, v \in \Sigma^+$ and $L_v \neq \emptyset$, we must have $L_v \cap L_{uv} = \emptyset$, and so $L_v \neq L_{uv}$. This means that, for any $t \in T_L$, if $1t \neq n$, then $1t \neq it$ for any $i \neq 1$. So $t \in G_n$, and $T_L \subseteq G_n$.

2. Since $T_L \subseteq G_n$, $\sigma(L) \leqslant |G_n| = g(n)$. Let $t \in G_n$ be any transformation. Note that $nt = n$ is fixed. There are two cases for $t$:

(a) $1t = n$: For any $i \neq 1, n$, there are $n-1$ choices for $it$: $2, 3, \ldots, n$. Thus there are $(n-1)^{n-2}$ such $t$'s.

(b) $1t \neq n$: There are $n-2$ choices for $1t$: $2, 3, \ldots, n-1$. For any $i \neq 1, n$, $it$ can be chosen from $\{2, 3, \ldots, n\} \setminus \{1t\}$. There are $(n-2)(n-2)^{n-2}$ such $t$'s.

Altogether, we have $g(n) = (n-1)^{n-2} + (n-2)^{n-1}$.

3. Assume that $T_L \subseteq G_n$, and let $f$ be the only accepting quotient. If $L$ is not suffix-free, then there exist nonempty words $u$ and $v$ such that $v, uv \in L$. Let $t_u$ and $t_v$ be the transformations by $u$ and $v$, and let $i = 1t_u$. Then $i \neq 1$, $f \neq n$, and $1t_v = f = 1t_{uv} = 1t_ut_v = it_v$, which contradicts the fact that $t_v \in G_n$. Therefore $L$ is suffix-free. □

Next, we construct a large semigroup that can be the syntactic semigroup of a suffix-free regular language. Let $C_k^n$ be the binomial coefficient, and let

$$P_n = \{t \in G_n \mid \text{ for all } i, j \in Q \text{ where } i \neq j, \text{ we have } it = jt = n \text{ or } it \neq jt\}.$$

**Proposition 7.** *For $n \geqslant 3$, $P_n$ is a subsemigroup of $G_n$, and its cardinality is*

$$p(n) = |P_n| = \sum_{k=1}^{n-1} C_k^{n-1}(n-1-k)!C_{n-1-k}^{n-2}.$$

*Proof.* We know that a transformation $t$ is in $P_n$ if and only if the following hold:

(a) $it \neq 1$ for all $i \in Q$, and $nt = n$;
(b) for all $i, j \in Q$, such that $i \neq j$, either $it = jt = n$ or $it \neq jt$.

Clearly $P_n \subseteq G_n$. For any $t_1, t_2 \in P_n$, consider the composition $t_1t_2$. Since $1 \notin \text{img } t_2$, then $1 \notin \text{img}(t_1t_2)$. We also have $nt_1t_2 = nt_2 = n$. Pick $i, j \in Q$ such that $i \neq j$. Suppose $it_1t_2 \neq n$ or $jt_1t_2 \neq n$. If $it_1t_2 = jt_1t_2$, then $it_1 = jt_1$ and thus $i = j$, a contradiction. Hence $t_1t_2 \in P_n$, and $P_n$ is a subsemigroup of $G_n$.

Let $t \in P_n$ be any transformation. Note that $nt = n$ is fixed. Let $Q' = Q \setminus \{n\}$, and $Q'' = Q \setminus \{1, n\}$. Suppose $k$ elements in $Q'$ are mapped to $n$ by $t$, where $0 \leqslant k \leqslant n-1$; then there are $C_k^{n-1}$ choices of these elements. For the set $D$ of the remaining $n-1-k$ elements, which must be mapped by $t$ to pairwise different elements of $Q''$, there are $C_{n-1-k}^{n-2}(n-1-k)!$ choices for the mapping $t|_D$. When $k = 0$, there is no such $t$ since $|Dt| = n-1 > n-2 = |Q''|$. Altogether, the cardinality of $P_n$ is $|P_n| = \sum_{k=1}^{n-1} C_k^{n-1}(n-1-k)!C_{n-1-k}^{n-2}$. □

We now construct a generating set of size $n$ for $P_n$, which can be used to show that there exist DFA's accepting suffix-free regular languages with quotient complexity $n$ and syntactic complexity $p(n)$.

**Proposition 8.** *When $n \geqslant 3$, the semigroup $P_n$ can be generated by the following set $I_n$ of transformations of $Q$: $I_3 = \{a, b\}$, where $a = [3, 2, 3]$ and $b = [2, 3, 3]$; $I_4 = \{a, b, c\}$, where $a = [4, 3, 2, 4]$, $b = [2, 4, 3, 4]$, $c = [2, 3, 4, 4]$; for $n \geqslant 5$, $I_n = \{a_0, \ldots, a_{n-1}\}$, where*

$$\begin{aligned} a_0 &= [\, n,\ 3,\ 2,\ 4,\ \ \ldots,\ n-1, n\,], \\ a_1 &= [\, n,\ 3,\ 4, \ldots, n-1,\ \ 2,\ \ n\,], \\ a_i &= [\, 2, \ldots, i,\ n,\ i+1,\ \ \ldots,\ n\,], \end{aligned}$$

*for $i = 2, \ldots, n-1$. That is, $a_0 = \binom{1}{n}(2,3)$, $a_1 = \binom{1}{n}(2, 3, \ldots, n-1)$, and $ja_i = j+1$ for $j = 1, \ldots, i-1$, $ia_i = n$, and $ja_i = j$ for $j = i+1, \ldots, n$.*

*Proof.* First note that $I_n$ is a subset of $P_n$, and so $\langle I_n \rangle$, the semigroup generated by $I_n$, is a subset of $P_n$. We now show that $P_n \subseteq \langle I_n \rangle$.

Let $t$ be any transformation in $P_n$. Note that $nt = n$ is fixed. Let $Q' = Q \setminus \{n\}$. Let $E_t = \{j \in Q' \mid jt = n\}$, $D_t = Q' \setminus E_t$, and $Q'' = Q \setminus \{1, n\}$. Then $D_t t \subseteq Q''$, and $|E_t| \geqslant 1$, since $|Q''| < |Q'|$. We prove by induction on $|E_t|$ that $t \in \langle I_n \rangle$.

First, note that $\langle a_0, a_1 \rangle$, the semigroup generated by $\{a_0, a_1\}$, is isomorphic to the symmetric group $\mathfrak{S}_{Q''}$ by Theorem 1. Consider $E_t = \{i\}$ for some $i \in Q'$. Then $ia_i = it = n$. Moreover, since $D_t a_i, D_t t \subseteq Q''$, there exists $\pi \in \langle a_0, a_1 \rangle$ such that $(ja_i)\pi = jt$ for all $j \in D_t$. Then $t = a_i \pi \in \langle I_n \rangle$.

Assume that any transformation $t \in P_n$ with $|E_t| < k$ can be generated by $I_n$, where $1 < k < n - 1$. Consider $t \in P_n$ with $|E_t| = k$. Suppose $E_t = \{e_1, \ldots, e_{k-1}, e_k\}$. Let $s \in P_n$ be such that $E_s = \{e_1, \ldots, e_{k-1}\}$. By assumption, $s$ can be generated by $I_n$. Let $i = e_k s$; then $i \in Q''$, and $e_j(sa_i) = n$ for all $1 \leqslant j \leqslant k$. Moreover, we have $D_t(sa_i) \subseteq Q''$. Thus, there exists $\pi \in \langle a_0, a_1 \rangle$ such that, for all $d \in D_t$, $d(sa_i\pi) = dt$. Altogether, for all $e_j \in E_t$, we have $e_j(sa_i\pi) = e_j t = n$, for all $d \in D_t$, $d(sa_i\pi) = dt$, and $n(sa_i\pi) = nt = n$. Thus $t = sa_i\pi$, and $t \in \langle I_n \rangle$.

Therefore $P_n = \langle I_n \rangle$. □

**Proposition 9.** *For $n \geqslant 5$, let $\mathcal{A}_n = \{Q, \Sigma, \delta, 1, F\}$ be the DFA with alphabet $\Sigma = \{a_0, a_1, \ldots, a_{n-1}\}$, where each $a_i$ defines a transformation as in Proposition 8, and $F = \{2\}$. Then $L = L(\mathcal{A}_n)$ has quotient complexity $\kappa(L) = n$, and syntactic complexity $\sigma(L) = p(n)$. Moreover, $L$ is suffix-free.*

*Proof.* First we show that all the states of $\mathcal{A}_n$ are reachable: 1 is the initial state, state $n$ is reached by $a_1$, and for $2 \leq i \leq n-1$, state $i$ is reached by $a_i^{i-1}$. Also, the initial state 1 accepts $a_2$ while state $i$ rejects $a_2$ for all $i \neq 1$. For $2 \leq i \leq n-1$, state $i$ accepts $a_1^{n-i}$, while state $j$ rejects it, for all $j \neq i$. Also $n$ is the empty state. Thus all the states of $\mathcal{A}_n$ are distinct, and $\kappa(L) = n$.

By Proposition 8, the syntactic semigroup of $L$ is $P_n$. The syntactic complexity of $L$ is $\sigma(L) = |P_n| = p(n)$. By Proposition 6, $L$ is suffix-free. □

As shown in Table 1 on p. 11, the size of $\Sigma$ cannot be decreased for $n \leqslant 5$.

By Proposition 6, the upper bound on the syntactic complexity of suffix-free regular languages is achieved by the largest subsemigroup of $G_n$. We conjecture that $P_n$ is such a subsemigroup.

*Conjecture 10 (Suffix-Free Regular Languages). If $L$ is a suffix-free regular language with $\kappa(L) = n \geqslant 2$, then $\sigma(L) \leqslant p(n)$ and this is a tight bound.*

We prove the conjecture for $n \leqslant 4$:

*Proof.* By Proposition 6, the syntactic semigroup $T_L$ of a suffix-free regular language $L$ is contained in $G_n$. For $n \in \{2, 3\}$, $p(n) = g(n)$. So $p(n)$ is an upper bound, and it is tight by Proposition 9. For $n = 4$, there are 17 transformations in $G_4$ and 13 in $P_4$. Transformations $r_1 = [3, 2, 2, 4]$ and $r_2 = [2, 3, 3, 4]$ in $G_4$ are such that $\langle r_i \rangle \not\subseteq G_4$ for $i = 1, 2$. Thus $T_L$ can contain neither $r_1$ nor $r_2$. Two other transformations, $s_1 = [4, 2, 2, 4]$ and $s_2 = [4, 3, 3, 4]$, in $G_4$ are such that $s_1$

conflicts with $t_1 = [3, 2, 4, 4] \in P_4$ ($t_1 s_1 = [2, 2, 4, 4] \notin G_4$), and $s_2$ conflicts with $t_2 = [2, 3, 4, 4]$ ($t_2 s_2 = [3, 3, 4, 4] \notin G_4$). Thus $\sigma(L) \leqslant 13$. By Proposition 9, the bound is tight. □

## 6 Bifix-Free Regular Languages

Let $L$ be regular and bifix-free with $\kappa(L) = n$. From Sections 4 and 5 we have:

1. $L$ has $\varepsilon$ as a quotient, and this is the only accepting quotient;
2. $L$ has $\emptyset$ as a quotient;
3. $L$ as a quotient is uniquely reachable.

Let $\mathcal{A}$ be the quotient automaton of $L$, with $Q = \{1, \ldots, n\}$ as the set of states. As in Section 5, we assume that $n-1$ corresponds to the quotient $\varepsilon$, and $n$ is the empty state. Consider the set

$$H_n = \{t \in G_n \mid (n-1)t = n\}.$$

Let $h(n) = |H_n|$. The following is an observation similar to Proposition 6.

**Proposition 11.** *If $L$ is a regular language with quotient complexity $n$ and syntactic semigroup $T_L$, then the following hold:*

1. *If $L$ is bifix-free, then $T_L$ is a subset of $H_n$.*
2. *If $L$ is bifix-free and $n \geqslant 3$, then $\sigma(L) \leqslant h(n) = (n-1)^{n-3} + (n-2)^{n-2}$.*
3. *If $\varepsilon$ is the only accepting quotient of $L$, and $T_L \subseteq H_n$, then $L$ is bifix-free.*

*Proof.* 1. Since $L$ is suffix-free, $T_L \subseteq G_n$. Since $L$ is also prefix-free, it has $\varepsilon$ and $\emptyset$ as quotients. By assumption, $n-1 \in Q$ corresponds to the quotient $\varepsilon$. Thus for any $t \in T_L$, $(n-1)t = n$, and so $T_L \subseteq H_n$.
2. To calculate the size of $H_n$, we analyze the two possible cases of $t \in H_n$:

(a) $1t = n$. For $i = 2, \ldots, n-2$, $it$ can be chosen from $\{2, \ldots, n\}$. There are $(n-1)^{n-3}$ such transformations.

(b) $1t \neq n$. For $i = 2, \ldots, n-2$, $it$ can be chosen from $\{2, \ldots, n\} \setminus \{1t\}$. There are $(n-2)(n-2)^{n-3} = (n-2)^{n-2}$ such transformations.

Altogether, we have $|T_L| \leqslant |H_n| = h(n) = (n-1)^{n-3} + (n-2)^{n-2}$.
3. Since $\varepsilon$ is the only accepting quotient of $L$, $L$ is prefix-free. Since $T_L \subseteq H_n \subseteq G_n$, $L$ is suffix-free by Proposition 6. Therefore $L$ is bifix-free. □

We now find a large semigroup that can be the syntactic semigroup of a bifix-free regular language. Let

$$R_n = \{t \in H_n \mid it = jt = n \text{ or } it \neq jt \text{ for all } 1 \leq i, j \leq n\}.$$

**Proposition 12.** *For $n \geqslant 3$, $R_n$ is a subsemigroup of $H_n$, and its cardinality is*

$$r(n) = |R_n| = \sum_{k=0}^{n-2} \left(C_k^{n-2}\right)^2 (n-2-k)!$$

*Proof.* First we show that $R_n$ is a semigroup. Let $t_1, t_2$ be any transformations in $R_n$. Since $1 \notin \text{img } t_2$, also $1 \notin \text{img}(t_1t_2)$. Since $nt_1 = nt_2 = n$, we have $nt_1t_2 = nt_2 = n$. Since $(n-1)t_1 = n$, also $(n-1)t_1t_2 = nt_2 = n$. Fix arbitrary $i, j \in Q$, where $i \neq j$. Note that $it_1t_2 \neq n$ implies $it_1 \neq n$. If $it_1t_2 = jt_1t_2$, since $t_2 \in R_n$, then $it_1 = jt_1 \neq n$; because $t_1 \in R_n$, we have $i = j$, a contradiction. Thus $it_1t_2 = jt_1t_2 = n$ or $it_1t_2 \neq jt_1t_2$. So $t_1t_2 \in R_n$, and $R_n$ is a subsemigroup of $H_n$.

Pick any $t \in R_n$. Note that $(n-1)t = n$ and $nt = n$ are fixed, and $1 \notin \text{img } t$. Let $Q' = Q \setminus \{n-1, n\}$, $E = \{i \in Q' \mid it = n\}$, and $D = Q' \setminus E$. Suppose $|E| = k$, where $0 \leq k \leq n-2$; then there are $C_k^{n-2}$ choices of $E$. Elements of $D$ are mapped to pairwise different elements of $Q \setminus \{1, n\}$; then there are $C_{n-2-k}^{n-2}(n-2-k)!$ different $t|_D$. Altogether, we have $|R_n| = \sum_{k=0}^{n-2} \left(C_k^{n-2}\right)^2 (n-2-k)!$ □

**Proposition 13.** *For $n \geqslant 3$, let $Q' = Q \setminus \{n-1, n\}$ and $Q'' = Q \setminus \{1, n\}$. Then the semigroup $R_n$ can be generated by $J_n = \{t \in R_n \mid Q't = Q''$ and $it \neq jt$ for all $i, j \in Q'\}$.*

*Proof.* We want to show that $R_n = \langle J_n \rangle$. Since $J_n \subseteq R_n$, we have $\langle J_n \rangle \subseteq R_n$. Let $t \in R_n$. By definition, $(n-1)t = nt = n$. Let $E_t = \{i \in Q' \mid it = n\}$. If $E_t = \emptyset$, then $t \in J_n$; otherwise, there exists $x \in Q''$ such that $x \notin \text{img } t$. We prove by induction on $|E_t|$ that $t \in \langle J_n \rangle$.

First note that, for all $t \in J_n$, $t|_{Q'}$ is an injective mapping from $Q'$ to $Q''$. Consider $E_t = \{i\}$ for some $i \in Q'$. Since $|E_t| = 1$, $\text{img } t \cup \{x\} = Q''$. Let $t_1, t_2 \in J_n$ be defined by

$$\begin{aligned} t_1 &= [\ 2,\ 3,\ \ldots,\ \ i,\quad n-1,\quad i+1,\ \ \ldots,\ \ n-2,\ \ n, n\ ], \\ t_2 &= [\ x,\ 1t,\ 2t,\ \ldots, (i-1)t, (i+1)t, \ldots, (n-2)t, n, n\ ]. \end{aligned}$$

That is, $jt_1 = j+1$ for $j = 1, \ldots, i-1$, $it_1 = n-1$, $jt_1 = j$ for $j = i+1, \ldots, n-2$, and $1t_2 = x$, $jt_2 = (j-1)t$ for $j = 2, \ldots, i$, $jt_2 = jt$ for $j = i+1, \ldots, n-2$. Then $t_1t_2 = t$, and $t \in \langle J_n \rangle$.

Assume that any transformation $t \in R_n$ with $|E_t| < k$ can be generated by $J_n$, where $1 < k < n-2$. Consider $t \in R_n$ with $|E_t| = k$. Suppose $E_t = \{e_1, \ldots, e_{k-1}, e_k\}$, and let $D_t = Q' \setminus E_t = \{d_1, \ldots, d_l\}$, where $l = n-2-k$. By assumption, all $s \in R_n$ with $|E_s| = k-1$ can be generated by $J_n$. Let $s$ be such that $E_s = \{1, \ldots, k-1\}$; then $1s = \cdots = (k-1)s = n$. In addition, let $ks = x$, and let $(k+j)s = d_jt$ for $j = 1, \ldots, l$. Let $t' \in J_n$ be such that $e_jt' = j$ for $j = 1, \ldots, k-1$, $kt' = n-1$, and $d_jt' = k+j$ for $j = 1, \ldots, l$. Then $t's = t$, and $t \in \langle J_n \rangle$. Therefore, $R_n = \langle J_n \rangle$. □

**Proposition 14.** *For $n \geqslant 3$, let $\mathcal{A}_n = \{Q, \Sigma, \delta, 1, F\}$ be the automaton with alphabet $\Sigma$ of size $(n-2)!$, where each $a \in \Sigma$ defines a distinct transformation $t_a \in J_n$, and $F = \{n-1\}$. Then $L = L(\mathcal{A}_n)$ has quotient complexity $\kappa(L) = n$, and syntactic complexity $\sigma(L) = r(n)$. Moreover, $L$ is bifix-free.*

*Proof.* We first show that all the states of $\mathcal{A}_n$ are reachable. Note that there exists $a \in \Sigma$ such that $t_a = [2, 3, \ldots, n-1, n, n] \in J_n$. State $1 \in Q$ is the

initial state, and $a^{i-1}$ reaches state $i \in Q$ for $i = 2, \ldots, n$. Furthermore, for $1 \leq i \leq n-1$, state $i$ accepts $a^{n-1-j}$, while for $j \neq i$, state $j$ rejects it. Also, $n$ is the empty state. Thus all the states of $\mathcal{A}_n$ are distinct, and $\kappa(L) = n$.

By Proposition 13, the syntactic semigroup of $L$ is $R_n$. Hence the syntactic complexity of $L$ is $\sigma(L) = r(n)$. By Proposition 11, $L$ is bifix-free. □

We know by Proposition 11 that the upper bound on the syntactic complexity of bifix-free regular languages is reached by the largest subsemigroup of $H_n$. We conjecture that $R_n$ is such a subsemigroup. Since $r(n) = h(n)$ for $n = 2$, 3, and 4, $r(n)$ is an upper bound, and it is tight by Proposition 14.

*Conjecture 15 (Bifix-Free Regular Languages). If $L$ is a bifix-free regular language with $\kappa(L) = n \geqslant 2$, then $\sigma(L) \leqslant r(n)$ and this is a tight bound.*

The conjecture holds for $n = 5$ as we now show:

*Proof.* For $n = 5$, we have $h(5) = |H_5| = 43$, and $r(5) = |R_5| = 34$. There are two transformations $s_1 = [2, 3, 3, 5, 5]$ and $s_2 = [3, 2, 2, 5, 5]$ in $H_5$ such that $\langle s_i \rangle \not\subseteq H_5$ for $i = 1, 2$. Thus $T_L$ cannot contain them, and we reduce the bound to $\sigma(L) \leqslant 41$. Let $U_5 = H_5 \setminus (R_5 \cup \{s_1, s_2\}) = \{\tau_1, \ldots, \tau_7\}$. We found for each $\tau_i$ a unique $t_i \in R_5$ such that the semigroup $\langle \tau_i, t_i \rangle$ is not a subset of $H_5$:

$$\begin{array}{ll}
\tau_1 = [2, 4, 4, 5, 5], & t_1 = [3, 4, 2, 5, 5]; \\
\tau_2 = [3, 4, 4, 5, 5], & t_2 = [3, 5, 2, 5, 5]; \\
\tau_3 = [4, 2, 2, 5, 5], & t_3 = [2, 4, 3, 5, 5]; \\
\tau_4 = [4, 3, 3, 5, 5], & t_4 = [2, 5, 3, 5, 5]; \\
\tau_5 = [5, 2, 2, 5, 5], & t_5 = [3, 2, 4, 5, 5]; \\
\tau_6 = [5, 3, 3, 5, 5], & t_6 = [2, 3, 4, 5, 5]; \\
\tau_7 = [5, 4, 4, 5, 5], & t_7 = [3, 2, 5, 5, 5].
\end{array}$$

Since $\langle \tau_i, t_i \rangle \subseteq T_L$, if both $\tau_i$ and $t_i$ are in $T_L$, then $T_L \not\subseteq H_5$, and $L$ is not bifix-free by Proposition 11. Thus, for $1 \leqslant i \leqslant 7$, at most one of $\tau_i$ and $t_i$ can

**Table 1.** Syntactic complexities of prefix-, suffix-, and bifix-free regular languages

| | $n = 2$ | $n = 3$ | $n = 4$ | $n = 5$ | $n = 6$ |
|---|---|---|---|---|---|
| $\lvert\Sigma\rvert = 1$ | **1** | **2** | **3** | **4** | **5** |
| $\lvert\Sigma\rvert = 2$ | $*$ | **3**/**3**/$*$ | **11**/**11**/**7** | **49**/**49**/**20** | ? |
| $\lvert\Sigma\rvert = 3$ | $*$ | $*$ | **14**/**13**/$*$ | **95**/**61**/**31** | ? |
| $\lvert\Sigma\rvert = 4$ | $*$ | $*$ | **16**/ $*$ /$*$ | **110**/**67**/**32** | ? |
| $\lvert\Sigma\rvert = 5$ | $*$ | $*$ | $*$ | **119**/73/**33** | ? |
| $\lvert\Sigma\rvert = 6$ | $*$ | $*$ | $*$ | **125**/ ? /**34** | ? /501/ ? |
| $\cdots$ | | | | | |
| $n^{n-2}/p(n)/r(n)$ | **1**/**1**/**1** | **3**/**3**/**2** | **16**/**13**/**7** | **125**/73/**34** | **1296**/501/209 |
| Suffix-free : $g(n)$ | 1 | 3 | 17 | 145 | 1,649 |
| Bifix-free : $h(n)$ | 1 | 2 | 7 | 43 | 381 |

appear in $T_L$, and $|T_L| \leqslant 34$. Since $|R_5| = 34$ and $R_5$ is a semigroup, we have $\sigma(L) \leqslant 34 = r(5)$ as the upper bound for $n = 5$. This bound is reached by automaton $\mathcal{A}_5$ in Proposition 14. □

## 7 Conclusions

Each cell of Table 1 shows the syntactic complexity bounds of prefix-, suffix-, and bifix-free regular languages, in that order, with a particular alphabet size. The figures in bold type are tight bounds verified by *GAP*. To compute the bounds for suffix- and bifix-free regular languages, we enumerate semigroups generated by elements of $G_n$ and $H_n$ that are subsemigroups of $G_n$ and $H_n$ respectively, and record the largest ones. By Propositions 6 and 11, we get the desired bounds from the enumeration. The asterisk $*$ indicates that the bound is already tight for a smaller alphabet. The last three rows include the tight upper bound $n^{n-2}$ for prefix-free regular languages, conjectured upper bounds $p(n)$ for suffix-free and $r(n)$ for bifix-free regular languages, and weaker upper bounds (not tight in general) $g(n)$ for suffix-free and $h(n)$ for bifix-free regular languages.

## References


1. Berstel, J., Perrin, D., Reutenauer, C.: Codes and Automata (Encyclopedia of Mathematics and its Applications). Cambridge University Press, Cambridge (2009)
2. Brzozowski, J.: Quotient complexity of regular languages. In: Dassow, J., Pighizzini, G., Truthe, B. (eds.) Proceedings of the 11th International Workshop on Descriptional Complexity of Formal Systems (DFS), Magdeburg, Germany, Otto-von-Guericke-Universität, pp. 25–42 (2009); to appear, J. Autom. Lang. Comb., Extended abstract at `http://arxiv.org/abs/0907.4547`
3. Brzozowski, J., Jirásková, G., Li, B., Smith, J.: Quotient complexity of bifix-, factor-, and subword-free regular languages. In: Proceedings of the 13th International Conference on Automata and Formal Languages, AFL (to appear, 2011), Full paper at `http://arxiv.org/abs/1006.4843v3`
4. Brzozowski, J., Ye, Y.: Syntactic complexity of ideal and closed languages. In: Mauri, G., Leporati, A. (eds.) Proceedings of the 15th International Conference on Developments in Language Theory (DLT). LNCS. Springer, Heidelberg (to appear, 2011), Full paper at `http://arxiv.org/abs/arXiv:1010.3263`
5. Ganyushkin, O., Mazorchuk, V.: Classical Finite Transformation Semigroups: An Introduction. Springer, Heidelberg (2009)
6. GAP-Group: GAP - Groups, Algorithms, Programming - a System for Computational Discrete Algebra (2010), `http://www.gap-system.org/`
7. Han, Y.S., Salomaa, K.: State complexity of basic operations on suffix-free regular languages. Theoret. Comput. Sci. 410(27-29), 2537–2548 (2009)
8. Han, Y.S., Salomaa, K., Wood, D.: Operational state complexity of prefix-free regular languages. In: Ésik, Z., Fülöp, Z. (eds.) Automata, Formal Languages, and Related Topics, Inst. of Informatics, pp. 99–115. University of Szeged, Hungary (2009)
9. Holzer, M., König, B.: On deterministic finite automata and syntactic monoid size. Theoret. Comput. Sci. 327(3), 319–347 (2004)

10. Hoyer, M.: Verallgemeinerung zweier sätze aus der theorie der substitutionengruppen. Math. Ann. 46, 539–544 (1895)
11. Krawetz, B., Lawrence, J., Shallit, J.: State complexity and the monoid of transformations of a finite set (2003), http://arxiv.org/abs/math/0306416v1
12. Maslov, A.N.: Estimates of the number of states of finite automata. Dokl. Akad. Nauk SSSR 194, 1266–1268 (1970) (Russian); English translation: Soviet Math. Dokl. 11, 1373–1375 (1970)
13. Myhill, J.: Finite automata and representation of events. Wright Air Development Center Technical Report, 57–624 (1957)
14. Nerode, A.: Linear automaton transformations. Proc. Amer. Math. Soc. 9, 541–544 (1958)
15. Piccard, S.: Sur les fonctions définies dans les ensembles finis quelconques. Fundamenta Mathematicae 24, 298–301 (1935)
16. Piccard, S.: Sur les bases du groupe symétrique et du groupe alternant. Commentarii Mathematici Helvetici 11, 1–8 (1938)
17. Pin, J.E.: Syntactic Semigroups. In: Handbook of Formal Languages. Word, Language, Grammar, vol. 1, Springer-Verlag New York, Inc., New York (1997)
18. Sierpiński, W.: Sur les suites infinies de fonctions définies dans les ensembles quelconques. Fund. Math. 24, 209–212 (1935)
19. Yu, S.: State complexity of regular languages. J. Autom. Lang. Comb. 6, 221–234 (2001)