Theory Comput Syst (2014) 54:277–292
DOI 10.1007/s00224-013-9515-7

Quotient Complexity of Closed Languages

Janusz Brzozowski · Galina Jirásková · Chenglong Zou

Published online: 19 November 2013

Abstract A language L is prefix-closed if, whenever a word w is in L, then every prefix of w is also in L. We define suffix-, factor-, and subword-closed languages in an analogous way, where by factor we mean contiguous subsequence, and by subword we mean scattered subsequence. We study the state complexity (which we prefer to call quotient complexity) of operations on prefix-, suffix-, factor-, and subword-closed languages. We find tight upper bounds on the complexity of the subword-closure of arbitrary languages, and on the complexity of boolean operations, concatenation, star, and reversal in each of the four classes of closed languages. We show that repeated applications of positive closure and complement to a closed language result in at most four distinct languages, while Kleene closure and complement give at most eight.

This research was supported by the Natural Sciences and Engineering Research Council of Canada grant OGP0000871, by VEGA grant 2/0183/11, and by grant APVV-0035-10.
A much shorter preliminary version of this work appeared as an arXiv preprint at http://arxiv.org/abs/0912.1034, and in F. Ablayev, E.W. Mayr, eds., Proc. 5th International Computer Science Symposium in Russia (CSR 2010). Volume 6072 of LNCS, Springer (2010), pp. 84–95.

Most of this work was done while C. Zou was at the David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada N2L 3G1.

J. Brzozowski (✉)
David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada N2L 3G1
e-mail: brzozo@uwaterloo.ca

G. Jirásková
Mathematical Institute, Slovak Academy of Sciences, Grešákova 6, 040 01 Košice, Slovakia
e-mail: jiraskov@saske.sk

C. Zou
Department of Mathematics, University of British Columbia, Vancouver, BC, Canada V6T 1Z4
e-mail: czou@math.ubc.ca

Keywords Closed language · Finite automaton · Quotient complexity · Regular language · State complexity

1 Introduction

The *state complexity of a regular language* L is the number of states in the minimal deterministic finite automaton (dfa) recognizing L. The *state complexity of an operation* in a subclass $\mathcal{C}$ of regular languages is defined as the worst-case size of the minimal dfa accepting the language resulting from the operation, considered as a function of the state complexities of the operands in $\mathcal{C}$.

The first results on the state complexity of reversal of a regular language are due to Mirkin [26] (1966), and of union, concatenation, and star of regular languages, to Maslov [25] (1970). For a general discussion of state complexity see the 2001 survey by Yu [31], the 2010 article by Brzozowski [6], and the reference lists in those papers. In 1994 the state complexities of concatenation, star, left and right quotients, reversal, intersection, and union of regular languages were examined in detail by Yu, Zhuang and K. Salomaa in [32]. The complexity of operations was also considered recently in several subclasses of regular languages: unary [28, 32], finite [12, 31], cofinite [3], prefix-free [20], suffix-free [19], bifix-, factor-, and subword-free [9], and ideal [8]. These studies show that the state complexity can be significantly lower in a subclass than in the general case. Here we examine state complexity in the classes of prefix-, suffix-, factor-, and subword-closed regular language; these classes are defined informally in the abstract and more formally in Sect. 3.

There are several reasons for considering closed languages. Subword-closed languages were studied in 1969 by Haines [18], in 1973 by Thierrin [29], and in 2010 by Okhotin [27]. Suffix-closed languages were considered in 1974 by Gill and Kou [15], in 1976 by Galil and Simon [14], in 1979 by Veloso and Gill [30], and in 2001 by Holzer, K. Salomaa, and Yu [21]. Factor-closed languages, also called *factorial*, have received some attention, for example, in 1990 by de Luca and Varricchio [13], and in 2005 by Avgustinovich and Frid [2]. The state complexities of the prefix-, suffix-, and factor-closure of a language were examined in 2009 by Kao, Rampersad, and Shallit [23]. Prefix-closed languages play a role in predictable semiautomata considered in 2009 by Brzozowski and Santean [10]. All four classes of closed languages were studied in 2009 by Ang and Brzozowski [1], and decision problems for closed languages were discussed in 2009 by Brzozowski, Shallit, and Xu [11]. Closed languages are closely related to ideals as follows [1]. A language is a *left ideal* (respectively, *right, two-sided*, *all-sided ideal*) if $L = \Sigma^* L$, (respectively, $L = L\Sigma^*$, $L = \Sigma^* L \Sigma^*$, and $L = \Sigma^* \sqcup\!\sqcup L$, where $\Sigma^* \sqcup\!\sqcup L$ is the shuffle of Σ^* with L). A non-empty language is a right (respectively left, two-sided, or all-sided) ideal if and only if its complement is a prefix-closed (respectively suffix-, factor-, or subword-closed) language. Closed languages are defined by the binary relation "is a prefix of" (respectively, "is a suffix of", "is a factor of", "is a subword of") [1], and are special cases of convex languages introduced in 1973 by Thierrin [29], and generalized in 2009 by Ang and Brzozowski [1]. Recent results concerning convex languages were surveyed in 2010 by Brzozowski [5]. The fact that the four classes of closed languages are related to each other permits us to obtain many results about them using similar methods.

The remainder of the paper is structured as follows. In Sect. 2 we discuss basic notions, including that of quotient complexity. Closure operations are studied in Sect. 3. The complexities of boolean operations, concatenation, star, and reversal are treated in Sect. 4. In Sect. 5 we examine the Kuratowski algebras generated by closed languages under the operations of complement and star, and complement and positive closure. Section 6 concludes the paper.

2 Quotient Complexity

The cardinality of a set S is denoted by $|S|$. If Σ is a non-empty finite *alphabet*, then Σ^* is the free monoid generated by Σ. A *word* is any element of Σ^*, and ε is the *empty word*. A *language* over Σ is any subset of Σ^*.

The following set operations are defined on languages: *complement* ($\overline{L} = \Sigma^* \setminus L$), *union* ($K \cup L$), *intersection* ($K \cap L$), *difference* ($K \setminus L$), and *symmetric difference* ($K \oplus L$). All four of these *boolean operations* with two arguments are denoted by $K \circ L$. We also define the *product* KL, usually called *concatenation* or *catenation:*

$$KL = \left\{w \in \Sigma^* \,\middle|\, w = uv, u \in K, v \in L\right\},$$

and the (Kleene) *star* L^* and the *positive closure* L^+:

$$L^* = \bigcup_{i \geq 0} L^i, \qquad L^+ = \bigcup_{i \geq 1} L^i.$$

The *reverse* w^R of a word w in Σ^* is defined as follows: $\varepsilon^R = \varepsilon$, and $(wa)^R = aw^R$. The *reverse* of a language L is denoted by L^R and is defined as $L^R = \{w^R \mid w \in L\}$.

Regular languages over an alphabet Σ are languages that can be obtained from the *set of basic languages* $\{\emptyset, \{\varepsilon\}\} \cup \{\{a\} \mid a \in \Sigma\}$, using a finite number of operations of union, product, and star. Such languages are usually denoted by regular expressions. If E is a regular expression, then $\mathcal{L}(E)$ is the language denoted by that expression. For example, $E = (\varepsilon \cup a)^* b$ denotes the language $\mathcal{L}(E) = (\{\varepsilon\} \cup \{a\})^*\{b\}$. We usually do not distinguish notationally between regular languages and regular expressions; the meaning is clear from the context.

A *deterministic finite automaton* (dfa) is a quintuple

$$\mathcal{D} = (Q, \Sigma, \delta, q_0, F),$$

where Q is a set of *states*, Σ is the *alphabet*, $\delta : Q \times \Sigma \to Q$ is the *transition function*, q_0 is the *initial state*, and F is the set of *final* or *accepting states*. The transition function of a dfa $\mathcal{D} = (Q, \Sigma, \delta, q_0, F)$ is extended to a function from $Q \times \Sigma^*$ to Q, and this extension is also denoted by δ. The *language accepted* by $\mathcal{D}$ is $L(\mathcal{D}) = \{w \mid \delta(q_0, w) \in F\}$, and $L(D_q)$ is the language accepted by the dfa $\mathcal{D}_q = (Q, \Sigma, \delta, q, F)$, which is the same as $\mathcal{D}$, except that its initial state is q; thus $L(\mathcal{D}) = L(\mathcal{D}_{q_0})$. If $L(\mathcal{D}_q)$ is empty, we call q the *empty state*, which is often called the *dead state*. States p and q of $\mathcal{D}$ are *equivalent* if $L(\mathcal{D}_p) = L(\mathcal{D}_q)$. If p and q are not equivalent, there must exist some word $w \in \Sigma^*$ which is in $L(\mathcal{D}_p)$ but not in $L(\mathcal{D}_q)$, or *vice versa*; then we say that p and q are *distinguishable* (by the word w).

A *nondeterministic finite automaton* (nfa)[1] is a quintuple

$$\mathcal{N} = (Q, \Sigma, \eta, Q_0, F),$$

where Q, Σ, and F are as in a dfa, $\eta : Q \times \Sigma \to 2^Q$ is the *transition function* and $Q_0 \subseteq Q$ is the *set of initial states*. If η also allows ε, that is, $\eta : Q \times (\Sigma \cup \{\varepsilon\}) \to 2^Q$, we call $\mathcal{N}$ an *ε-nfa*. The extended transition function $\hat{\eta} : 2^Q \times \Sigma^* \to 2^Q$ is defined as follows. Let $S \subseteq Q$ be a set of states of $\mathcal{N}$, and let $\eta_\varepsilon(S)$ be the set of states that can be reached from any state in S through ε-transitions. Then

$$\hat{\eta}(S, \varepsilon) = \eta_\varepsilon(S),$$

$$\hat{\eta}(S, xa) = \bigcup_{q \in \hat{\eta}(S,x)} \eta_\varepsilon\big(\eta(q, a)\big),$$

for all $a \in \Sigma$ and $x \in \Sigma^*$. We usually refer to $\hat{\eta}$ as η. The nfa or ε-nfa $\mathcal{N}$ accepts $x \in \Sigma^*$ if $\hat{\eta}(Q_0, x) \cap F \neq \emptyset$, and $L(\mathcal{N})$ is the set of all words accepted by $\mathcal{N}$. Thus any nfa or ε-nfa is equivalent to some dfa $\mathcal{D} = (Q', \Sigma, \delta, q'_0, F')$ in which $Q' = 2^Q$ is the set of states, δ is the transition function defined by $\delta(S, a) = \hat{\eta}(S, a)$ for every $S \in Q'$, $q'_0 = \eta_\varepsilon(Q_0)$ is the initial state, and $F' = \{S \subseteq Q \mid S \cap F \neq \emptyset\}$ is the set of final states. The language $L(\mathcal{N}_q)$ of a state q of nfa $\mathcal{N}$ is the language accepted by the nfa $\mathcal{N}_q = (Q, \Sigma, \eta, \eta_\varepsilon(\{q\}), F)$.

Our approach to quotient complexity follows closely that of [6]. Since state complexity is a property of a language, we prefer to define it in language-theoretic terms. The *left quotient*, or simply *quotient,* of a language L by a word w is the language $L_w = \{x \in \Sigma^* \mid wx \in L\}$. The *quotient complexity* of L is the number of distinct quotients of L and is denoted by $\kappa(L)$.

Quotients of regular languages [4, 6] can be computed as follows. First, the *ε-function* L^ε of a regular language L is

$$L^\varepsilon = \begin{cases} \emptyset, & \text{if } \varepsilon \notin L; \\ \varepsilon, & \text{if } \varepsilon \in L. \end{cases} \tag{1}$$

The quotient by a letter a in Σ is computed by induction:

$$b_a = \begin{cases} \emptyset, & \text{if } b \in \{\emptyset, \varepsilon\}, \quad \text{or } b \in \Sigma \text{ and } b \neq a; \\ \varepsilon, & \text{if } b = a. \end{cases} \tag{2}$$

$$\begin{aligned} &(\overline{L})_a = \overline{L_a}; \qquad (K \cup L)_a = K_a \cup L_a; \qquad (KL)_a = K_a L \cup K^\varepsilon L_a; \\ &\big(K^*\big)_a = K_a K^*. \end{aligned} \tag{3}$$

The quotient by a word $w \in \Sigma^*$ is computed by induction on the length of w:

$$L_\varepsilon = L; \qquad L_{wa} = (L_w)_a. \tag{4}$$

[1] In contrast to some authors, we use a *set* of initial states, since we require the reverse of an nfa to be an nfa.

A quotient L_w is *accepting* if $\varepsilon \in L_w$; otherwise it is *rejecting*. If the empty language is one of the quotients of a language L, then we say that L *has the empty quotient*.

It is well known that there is a one-to-one correspondence between the quotients of a regular language L and the states of the minimal dfa accepting L. Hence the quotient complexity of L is equal to the state complexity of L. Sometimes there are some advantages to using quotient complexity [6]; in other cases, it may be preferable to use state complexity and automata.

The formulas given next can be used to establish upper bounds on quotient complexity. To simplify the notation, we write $(L_w)^\varepsilon$ as L_w^ε.

Proposition 1 ([4, 6]) *If K and L are regular languages over an alphabet Σ, and u and v below are in Σ^+, then*

$$(\overline{L})_w = \overline{L_w}; \qquad (K \circ L)_w = K_w \circ L_w; \tag{5}$$

$$(KL)_w = K_w L \cup K^\varepsilon L_w \cup \left(\bigcup_{w=uv} K_u^\varepsilon L_v \right); \tag{6}$$

$$\left(L^*\right)_\varepsilon = \varepsilon \cup LL^*, \qquad \left(L^*\right)_w = \left(L_w \cup \bigcup_{w=uv} \left(L^*\right)_u^\varepsilon L_v \right) L^* \quad \textit{for } w \in \Sigma^+. \tag{7}$$

3 Closure Operations

If $w = uxv$ for some u, v, x in Σ^*, then u is a *prefix* of w, v is a *suffix* of w, and x is a *factor* of w. If $w = w_0 a_1 w_1 \cdots a_n w_n$, where $a_1, \dots, a_n \in \Sigma$, and $w_0, \dots, w_n \in \Sigma^*$, then the word $a_1 \cdots a_n$ 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 a similar way, we define *suffix-*, *factor-*, and *subword-closed* languages. A language is *closed* if it is prefix-, suffix-, factor-, or subword-closed.

Let $\trianglelefteq$ be a partial order on Σ^*; the $\trianglelefteq$*-closure* of a language L is the language

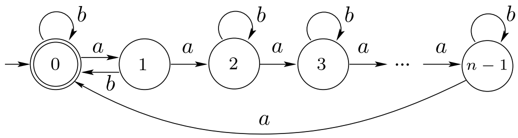
$$_{\trianglelefteq}L = \left\{ x \in \Sigma^* \,\middle|\, x \trianglelefteq w \text{ for some } w \in L \right\}.$$

For our applications, the partial order becomes one of the relations "is a prefix of", "is a suffix of", "is a factor of", or "is a subword of".

The worst-case quotient complexity for closure was studied by Kao, Rampersad, and Shallit [23]. For suffix-closure, the bound $2^n - 1$ holds in the case where L does not have the empty quotient. We add the case where L has the empty quotient; here the bound is 2^{n-1}. Subword-closure was previously studied by Gruber, Holzer and Kutrib [16, 17] and Okhotin [27], but tight upper bounds were not established; our next theorem solves this problem. For completeness, we provide the proofs for all four closure operations.

Theorem 1 (Closure Operations) *Let $n \geq 2$. Let L be a regular language over an alphabet Σ with $\kappa(L) = n$.*

Fig. 1 An n-state dfa of a language L that does not have the empty quotient



1. *If K is the prefix-closure of L, then $\kappa(K) \le n$, and the bound is tight if $|\Sigma| \ge 1$.*
2. *If K is the suffix-closure of L, then $\kappa(K) \le 2^n - 1$ if L does not have empty quotient and $\kappa(K) \le 2^{n-1}$ otherwise; both bounds are tight if $|\Sigma| \ge 2$.*
3. *If K is the factor-closure of L, then $\kappa(K) \le 2^{n-1}$, and the bound is tight if $|\Sigma| \ge 2$.*
4. *If K is the subword-closure of L, then $\kappa(K) \le 2^{n-2} + 1$, and the bound is tight if $|\Sigma| \ge n - 2$.*

Proof We assume that the given regular language L is represented by its minimal dfa $\mathcal{D} = (Q, \Sigma, \delta, q_0, F)$, and that the quotient complexity of L is n.

1. To get the dfa for the prefix-closure of L, we only need to make each non-empty state of $\mathcal{D}$ accepting. This gives the upper bound n. For tightness, consider the unary language $L = \{a^i \mid i \le n - 2\}$. The prefix-closure of L is the same language. Thus the upper bound is tight if $|\Sigma| \ge 1$.

2. We can construct an nfa for the suffix-closure of L by making each non-empty state of $\mathcal{D}$ initial. Then we apply the subset construction to this nfa.

If L does not have the empty quotient, then the construction begins with the set Q, which is non-empty. Since $\mathcal{D}$ is deterministic, each set reached by a letter from Σ from a non-empty set is non-empty. Hence the empty set cannot be reached, and so the subset dfa has at most $2^n - 1$ states.

If L has the empty quotient, then the nfa for the suffix-closure of L has $n - 1$ states, and the minimal equivalent dfa has at most 2^{n-1} states.

To prove tightness in the case where L does not have the empty quotient, consider the language L defined by the n-state dfa shown in Fig. 1.

Construct an nfa for the suffix-closure of L from this minimal dfa by making all states initial. Since the word a^{n-i} is accepted by the nfa only from state i for $i = 0, 1, \ldots, n - 1$, no two different states of the corresponding subset dfa are equivalent.

Let us show that the corresponding subset dfa has $2^n - 1$ reachable states. The proof is by induction on the size of subsets, going from n down to 1. The basis, $|S| = n$, holds since $\{0, 1, \ldots, n - 1\}$ is the initial state. Assume that each set of size k is reachable and let S be a set of size $k - 1$. If S contains state 0 but does not contain state 1, then it can be reached from the set $S \cup \{1\}$ of size k by b. If S contains both 0 and 1, then there is a state i such that $i \in S$ and $i + 1 \notin S$. Then S can be reached from $\{s - i \bmod n \mid s \in S\}$ by a^i. The latter set contains 0 and does not contain 1, and so is reachable. If a non-empty S does not contain 0, then it can be reached from $\{s - \min S \mid s \in S\}$, which contains 0, by $a^{\min S}$. Hence the subset dfa has $2^n - 1$ reachable states, and no two different states are equivalent.

Now consider the case where a language has the empty quotient. Let L be the language defined by the n-state dfa shown in Fig. 2. Remove state $n - 1$ and the transitions going to it, and then construct an nfa by making all non-empty states

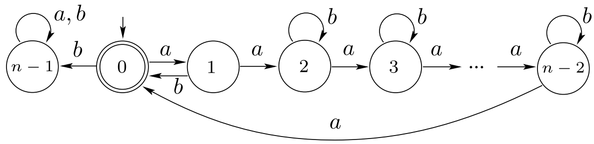


Fig. 2 An n-state dfa of a language L that has the empty quotient

initial. Since the word $(ab)^n$ is accepted by the nfa only from state 0, and the word $a^{n-1-i}(ab)^n$ only from state i for $i = 1, 2, \dots, n-2$, no two different states of the corresponding subset dfa are equivalent. The proof of reachability of all non-empty subsets of $\{0, 1, \dots, n-2\}$ is the same as above, and the empty set is reached from $\{0\}$ by b. Hence the subset dfa has 2^{n-2} reachable and pairwise distinguishable states.

3. To find an nfa for the factor-closure of L, we make all the non-empty states of the minimal automaton of L both accepting and initial, and delete the empty state, if it exists. If there is no empty state, then the factor-closure of L is Σ^*. Otherwise, the nfa for the factor-closure has at most 2^{n-1} states. The language L defined by the minimal automaton of Fig. 2 meets this bound.

4. If L does not have the empty quotient, then its subword-closure is Σ^*. Otherwise, to get an ε-nfa for the subword-closure of L, we remove the empty state of $\mathcal{D}$, and add an ε-transition from state p to state q whenever there is a transition from p to q in $\mathcal{D}$. Since the initial state of the ε-nfa can reach every non-empty state by ε-transitions, no other subset containing the initial state can be reached. Hence there are at most $2^{n-2}+1$ reachable subsets.

To prove tightness, if $n = 2$, let $\Sigma = \{a, b\}$; then $L = a^*$ meets the bound. If $n \geq 3$, let $\Sigma = \{a_1, \dots, a_{n-2}\}$, and

$$L = \bigcup_{a_i \in \Sigma} a_i \bigl(\Sigma \setminus \{a_i\}\bigr)^*.$$

Thus L consists of all the words over Σ in which the first letter of the word occurs exactly once. Now consider any subword x obtained by deleting some letters from a word w in $a_i(\Sigma \setminus \{a_i\})^*$. If a_i is deleted from w, then a_i does not appear in x. If another letter, a_j, is deleted from w, but another occurrence of a_j remains in x, then x is still in L, and need not be taken into account. Consequently, if K is the subword-closure of L, then

$$K = L \cup \bigl\{w \in \Sigma^* \bigm| \text{at least one letter of } \Sigma \text{ is missing in } w\bigr\}.$$

For each boolean vector $b = (b_1, b_2, \dots, b_{n-2})$, we now define the word $w(b) = w_1 w_2 \cdots w_{n-2}$, in which $w_i = \varepsilon$ if $b_i = 0$ and $w_i = a_i$ if $b_i = 1$. Consider ε, and each word $a_1 w(b)$. All the quotients of K by these $2^{n-2}+1$ words are distinct: For each binary vector b, we have $a_1 a_2 \cdots a_{n-2} \in K_\varepsilon \setminus K_{a_1 w(b)}$. Let b and b' be two different vectors with $b_i = 0$ and $b'_i = 1$. Then we have $a_1 a_2 \cdots a_{i-1} a_{i+1} a_{i+2} \cdots a_{n-2} \in K_{a_1 w(b)} \setminus K_{a_1 w(b')}$. Thus all the quotients are distinct and $\kappa(K) = 2^{n-2}+1$. □

Example 1 Let $L = a(b \cup c)^* \cup b(a \cup c)^* \cup c(a \cup b)^*$. Then L has five distinct quotients:

$$\begin{aligned}
L^{\varepsilon} &= L, \\
L_a &= L_{ab} = L_{ac} = (b \cup c)^*, \\
L_b &= L_{ba} = L_{bc} = (a \cup c)^*, \\
L_c &= L_{ca} = L_{cb} = (a \cup b)^*, \\
L_{aa} &= L_{aaa} = L_{aab} = L_{aac} = L_{bb} = L_{cc} = \emptyset.
\end{aligned}$$

The subword-closure K of L has the form

$$K = L \cup (b \cup c)^* \cup (a \cup c)^* \cup (a \cup b)^*,$$

and K has nine distinct quotients:

$$\begin{aligned}
K_{\varepsilon} &= K, \\
K_a &= K_b = K_c = (b \cup c)^* \cup (a \cup c)^* \cup (a \cup b)^*, \\
K_{aa} &= (a \cup c)^* \cup (a \cup b)^*, \\
K_{ab} &= (b \cup c)^* \cup (a \cup b)^*, \\
K_{ac} &= (b \cup c)^* \cup (a \cup c)^*, \\
K_{aab} &= (a \cup b)^*, \\
K_{aac} &= (a \cup c)^*, \\
K_{abc} &= (b \cup c)^*, \\
K_{aabc} &= \emptyset.
\end{aligned}$$

4 Basic Operations on Closed Languages

Now we study the quotient complexity of operations on closed languages. For regular languages, the following tight upper bounds are known [25, 26, 32]: mn for boolean operations, $m2^n - 2^{n-1}$ for product, $2^{n-1} + 2^{n-2}$ for star, and 2^n for reversal. The bounds for closed languages are smaller in most cases.

Theorem 2 (Boolean Operations) *Let K and L be prefix-closed* (*or factor-closed*, *or subword-closed*) *languages over an alphabet Σ with $\kappa(K) = m$ and $\kappa(L) = n$. Then*

1. $\kappa(K \cap L) \le mn - (m + n - 2)$,
2. $\kappa(K \cup L), \kappa(K \oplus L) \le mn$,
3. $\kappa(K \setminus L) \le mn - (n - 1)$.

For suffix-closed languages, $\kappa(K \circ L) \le mn$. *All bounds are tight if $|\Sigma| \ge 2$, except for the union and difference of suffix-closed languages, where we assume $|\Sigma| \ge 4$.*

Proof Recall that the complement of a prefix-closed (respectively, suffix-closed, factor-closed, or subword-closed) language is a right ideal (respectively, left, two-sided, or all-sided ideal). By De Morgan's laws and the results from [8, Theorem 7, p. 45], we have

$$\kappa(K \cap L) = \kappa(\overline{\overline{K} \cup \overline{L}}) = \kappa(\overline{K} \cup \overline{L}) \le mn - (m+n-2),$$

$$\kappa(K \cup L) = \kappa(\overline{\overline{K} \cap \overline{L}}) = \kappa(\overline{K} \cap \overline{L}) \le mn,$$

$$\kappa(K \setminus L) = \kappa(K \cap \overline{L}) = \kappa(\overline{L} \setminus \overline{K}) \le mn - (n-1),$$

$$\kappa(K \oplus L) = \kappa\big((K \setminus L) \cup (L \setminus K)\big) = \kappa\big((\overline{L} \setminus \overline{K})\big) \cup (\overline{K} \setminus \overline{L}) = \kappa(\overline{K} \oplus \overline{L}) \le mn.$$

□

Remark 1 If L is prefix-closed, then either $L = \Sigma^*$ or L has the empty quotient. Moreover, each quotient of L is either accepting or empty.

Remark 2 For a suffix-closed language L, if v is a suffix of w, then $L_w \subseteq L_v$. In particular, $L_w \subseteq L_\varepsilon = L$ for each word w in Σ^*.

Theorem 3 (Product) *Let $m, n \ge 2$. Let K and L be closed languages over an alphabet Σ with $\kappa(K) = m$, $\kappa(L) = n$, and let k be the number of accepting quotients of K.*

1. *If K and L are prefix-closed, then $\kappa(KL) \le (m+1) \cdot 2^{n-2}$.*
2. *If K and L are suffix-closed, then $\kappa(KL) \le (m-k)n + k$.*
3. *If K and L are both factor-closed or both subword-closed, then $\kappa(KL) \le m + n - 1$.*

The first two bounds are tight if $|\Sigma| \ge 3$, and the third, if $|\Sigma| \ge 2$. If $\kappa(K) = 1$ or $\kappa(L) = 1$, then $\kappa(KL) = 1$.

Proof If $m = 1$, then $K = \emptyset$ or $K = \Sigma^*$. Hence $KL = \emptyset$ or $KL = \Sigma^*$, for if $L \ne \emptyset$, then $\varepsilon \in L$. Thus $\kappa(KL) = 1$. The case $n = 1$ is similar. Now let $m, n \ge 2$.

1. If K and L are prefix-closed, then $\varepsilon \in K$ and by Remark 1 both languages have the empty quotient. The quotient $(KL)_w$ is given by Eq. (6). If K_w is accepting, then L is always in the union, and there are 2^{n-2} non-empty subsets of non-empty quotients of L that can be added. Since there are $m-1$ accepting quotients of K, there are $(m-1)2^{n-2}$ such quotients of KL. If K_w is rejecting, then 2^{n-1} subsets of non-empty quotients of L can be added. Altogether $\kappa(KL) \le 2^{n-1} + (m-1)2^{n-2} = (m+1)2^{n-2}$.

For tightness, consider the prefix-closed languages K and L defined by the dfa's shown in Fig. 3, except in the case where $n = 2$, in which case let $L = \{a, c\}^*$.

Construct an ε-nfa for the language KL from these minimal dfa's by adding an ε-transition from states $q_0, q_1, \dots, q_{m-2}$ to state 0. The initial state of the ε-nfa is q_0, and the accepting states are $0, 1, \dots, n-2$. We show that there are $(m+1) \cdot 2^{n-2}$ reachable and pairwise distinguishable states in the corresponding subset dfa.

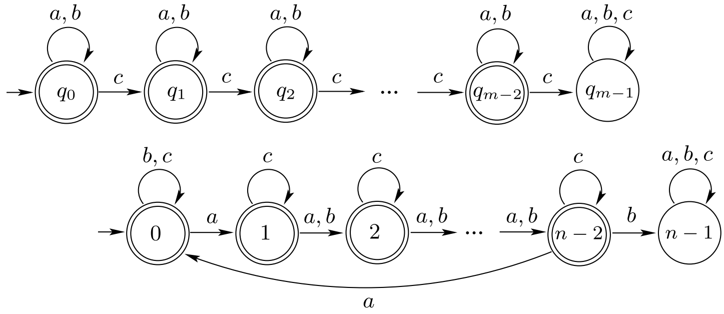


Fig. 3 Dfa's of prefix-closed languages K (*top*) and L (*bottom*)

State $\{q_0, 0\}$ is the initial state, and each state $\{q_0, 0, i_1, i_2, \dots, i_k\}$, where $1 \le i_1 < i_2 < \cdots < i_k \le n-2$, is reached from state $\{q_0, 0, i_2 - i_1, \dots, i_k - i_1\}$ by $ab^{i_1 - 1}$. For each subset S of $\{0, 1, \dots, n-2\}$ containing 0, each state $\{q_i\} \cup S$ with $1 \le i \le m-1$ is reached from $\{q_0\} \cup S$ by c^i. If a non-empty S does not contain 0, then $\{q_{m-1}\} \cup S$ is reached from $\{q_{m-1}\} \cup \{s - \min S \mid s \in S\}$, which contains 0, by $a^{\min S}$. State $\{q_{m-1}, n-1\}$ is reached from $\{q_{m-1}, n-2\}$ by word b.

To prove that no two states of the subset dfa are equivalent, notice that the word b^n is accepted by the minimal dfa for L only from state 0, and the word $a^{n-1-i}b^n$ only from state i $(1 \le i \le n-2)$. Therefore two different states $\{q_{m-1}\} \cup S$ and $\{q_{m-1}\} \cup T$ are distinguishable since S and T must differ by at least one state of the dfa for L, and state q_{m-1} is the empty state in the minimal dfa for K. It follows that states $\{q_i\} \cup S$ and $\{q_i\} \cup T$ are distinguishable as well since they go to two distinguishable states by c^{m-1-i}. States $\{q_i\} \cup S$ and $\{q_j\} \cup T$ with $i < j$ can be distinguished by $c^{m-1-j}b^n ab^n$. Hence the subset dfa has $(m+1) \cdot 2^{n-2}$ reachable and pairwise distinguishable states, and so $\kappa(KL) = (m+1)2^{n-2}$.

2. If K and L are suffix-closed, then, by Remark 2, for each word w in Σ^* and for all u, v in Σ^+, we have

$$(KL)_w = K_w L \cup K^\varepsilon L_w \cup \left(\bigcup_{w=uv} K_u^\varepsilon L_v \right) = K_w L \cup L_x$$

for some suffix x of w. If K_w is a rejecting quotient, there are at most $(m-k)n$ such quotients. If K_w is accepting, then $\varepsilon \in K_w$, and since $L_x \subseteq L_\varepsilon = L \subseteq K_w L$, we have $(KL)_w = K_w L$. There are at most k such quotients. Therefore there are at most $(m-k)n + k$ quotients in total.

To prove tightness, let K and L be the ternary suffix-closed languages defined by the dfa's of Fig. 4.

Consider the words $\varepsilon = a^0 b^0$, and $a^i b^j$ with $1 \le i \le m-1$ and $0 \le j \le n-1$. Let us show that all the quotients of KL by these words are distinct. Let $(i, j) \ne (k, \ell)$, and let $x = a^i b^j$ and $y = a^k b^\ell$. If $i < k$, take $z = a^{m-1-k} b^n c$. Then xz is in KL, while yz is not, and so $z \in (KL)_x \setminus (KL)_y$. If $i = k$ and $j < \ell$, take $z = a^m b^{n-1-\ell} c$. We again have $z \in (KL)_x \setminus (KL)_y$.

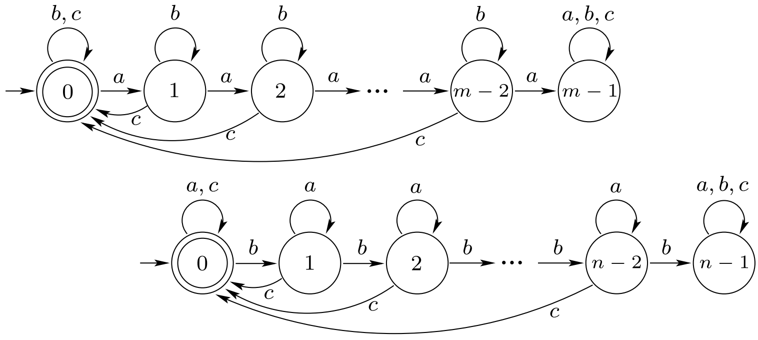


Fig. 4 Dfa's of suffix-closed languages K (*top*) and L (*bottom*)

Notice that, if the quotients K_{a^i} with $0 \le i \le k-1$ are accepting, then the resulting product has quotient complexity $(m-k)n+k$.

3. It suffices to derive the bound for factor-closed languages, because every subword-closed language is also factor-closed. Since factor-closed languages are suffix-closed, $\kappa(KL) \le (m-k)n+k$. Because K is prefix-closed, it has at most one rejecting quotient. Thus, $k = m-1$ and $\kappa(KL) \le m+n-1$.

To prove tightness, let

$$K = \{w \in \{a,b\}^* \mid a^{m-1} \text{ is not a subword of } w\}$$

and

$$L = \{w \in \{a,b\}^* \mid b^{n-1} \text{ is not a subword of } w\}.$$

Then K and L are subword-closed and $\kappa(K) = m$ and $\kappa(L) = n$. Consider the word $w = a^{m-1}b^{n-1}$. This word is not in the product KL. However, removing any non-empty subword from w results in a word in KL. Therefore, $\kappa(KL) \ge m+n-1$. □

Theorem 4 (Star) *Let $n \ge 3$, and let L be a closed language over an alphabet Σ with $\kappa(L) = n$.*

1. *If L is prefix-closed, then $\kappa(L^*) \le 2^{n-2}+1$.*
2. *If L is suffix-closed, then $\kappa(L^*) = n$ if $L = L^*$, and $\kappa(L^*) \le n-1$ if $L \ne L^*$.*
3. *If L is factor- or subword-closed, then $\kappa(L^*) \le 2$.*

The first bound is tight if $|\Sigma| \ge 3$, and all the other bounds are tight if $|\Sigma| \ge 2$. If $n = 1$, then $\kappa(L^*) \le 2$. If $n = 2$, then $\kappa(L^*) = 2$.

Proof Suppose L is closed under one of the four binary relations. If $n = 1$, then L is either empty or Σ^*, and L^* is either $\{\varepsilon\}$ or Σ^*. Thus $\kappa(L^*) \le 2$. If $n = 2$, then L cannot be empty, and must contain ε since it is closed. Hence L must be $\{\varepsilon\}$, since the quotient which is not L must be rejecting. Thus $L^* = \{\varepsilon\}$, and $\kappa(L^*) = 2$.

1. For every non-empty word w, the quotient $(L^*)_w$ is given by Eq. (7). If L is prefix-closed, then so is L^* and $(L^*)_w$. Thus, if $(L^*)_w$ is non-empty, then it contains ε.

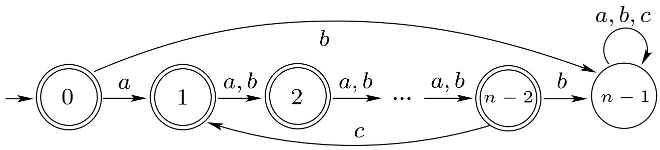


Fig. 5 An n-state dfa of prefix-closed language L

Hence $(L^*)_w \supseteq L^* \supseteq L$. Since $\emptyset$ and L are always contained in every non-empty quotient of L^*, there are at most 2^{n-2} non-empty quotients of L^*. Since there is at most one empty quotient, there are at most $2^{n-2}+1$ quotients in total. The quotient $(L^*)_\varepsilon$ has already been counted, since L is closed and $\varepsilon \in L$ implies $(L^*)_\varepsilon = LL^*$, which has the form of Eq. (7) for w in Σ^+.

Now let $n \geq 3$ and let L be the prefix-closed language defined by the dfa of Fig. 5; transitions not depicted in the figure go to state $n-1$.

Construct an ε-nfa for L^* by removing state $n-1$ and adding an ε-transition from all the remaining states to the initial state. Let us show that $2^{n-2}+1$ states are reachable and pairwise distinguishable in the corresponding subset dfa.

We first prove that each subset of $\{0, 1, \dots, n-2\}$ containing state 0 is reachable in the subset dfa. The proof is by induction on the size of the subsets. The basis, $|S| = 1$, holds since $\{0\}$ is the initial state of the subset dfa. Assume that each set of size k containing 0 is reachable, and let $S = \{0, i_1, i_2, \dots, i_k\}$, where $0 < i_1 < i_2 < \cdots < i_k \leq n-2$, be a set of size $k+1$. Then S is reached from the set $\{0, i_2 - i_1, \dots, i_k - i_1\}$ of size k by ab^{i_1-1}. Since the latter set is reachable by the induction hypothesis, the set S is reachable as well. The empty set can be reached from $\{0\}$ by b, and we have $2^{n-2}+1$ reachable states. To prove distinguishability, notice that b^{n-3} is accepted by the nfa only from state 1, and each word $b^{n-2-i}cb^{n-3}$ $(2 \leq i \leq n-2)$, only from state i.

2. If $L = L^*$, then $\kappa(L^*) = n$. Let $L = (a \cup ba^{n-2})^*$; then L is suffix-closed, $\kappa(L) = n$, and $L^* = L$.

Now suppose that $L \neq L^*$. For a non-empty suffix-closed language L, the quotient $(L^*)_\varepsilon$ is LL^*, which is of the same form as the quotients by a non-empty word w in Eq. (7). By that equation, we have

$$\left(L^*\right)_w = (L_w \cup L_{v_1} \cup \cdots \cup L_{v_k})L^*,$$

where the v_i are suffixes of w, and v_k is the shortest. By Remark 2, $(L^*)_w = L_{v_k}L^*$ for all $w \in \Sigma^*$, and $\kappa(L^*) \leq n$.

Assume for the sake of contradiction that $\kappa(L) = \kappa(L^*)$. Then we must have $(L^*)_x = (L^*)_y$ if and only if $L_x = L_y$. Since $L \neq L^*$, there exist $x, y \in L$ such that $xy \notin L$. Hence $L_x \neq L_\varepsilon$, since $y \in L_\varepsilon$ and $y \notin L_x$. But, by Eq. (7), $L^* \subseteq L_xL^* \subseteq (L^*)_x \subseteq L^*$ since $\varepsilon \in L_x$. So $(L)^*_x = (L)^*_\varepsilon$, which is a contradiction. Hence $\kappa(L) > \kappa(L^*)$ and $\kappa(L^*) \leq n-1$.

If $L = \varepsilon \cup \bigcup_{i=0}^{n-3} a^i b$, then L is suffix-closed, $\kappa(L) = n$, $L^* = (\bigcup_{i=0}^{n-3} a^i b)^*$, and $\kappa(L^*) = n-1$.

3. If each letter of Σ appears in some word of a factor-closed language L, then $L^* = \Sigma^*$ and $\kappa(L^*) = 1$. Otherwise, $\kappa(L^*) = 2$. The bound is met by the subword-closed language $L = \{w \in \{a,b\}^* \mid w = a^i \text{ and } 0 \le i \le n-2\}$. □

Since the operation of reversal commutes with complementation, the next theorem follows from the results on ideal languages [8, Theorem 11, p. 48].

Theorem 5 (Reversal) *Let $n \ge 2$. Let L be a closed language over an alphabet Σ with $\kappa(L) = n$.*

1. *If L is prefix-closed, then $\kappa(L^R) \le 2^{n-1}$, and the bound is tight if $|\Sigma| \ge 2$.*
2. *If L is suffix-closed, then $\kappa(L^R) \le 2^{n-1} + 1$, and the bound is tight if $|\Sigma| \ge 3$.*
3. *If L is factor-closed, then $\kappa(L^R) \le 2^{n-2} + 1$, and the bound is tight if $|\Sigma| \ge 3$.*
4. *If L is subword-closed, then $\kappa(L^R) \le 2^{n-2} + 1$, and the bound is tight if $|\Sigma| \ge 2n$.*

If $\kappa(L) = 1$, then $\kappa(L^R) = 1$.

Unary Languages Unary languages have special properties because the product of unary languages is commutative. The classes of prefix-closed, suffix-closed, factor-closed, and subword-closed unary languages all coincide. If a unary closed language L is finite, then either it is empty and so $\kappa(L) = 1$, or has the form $\{a^i \mid i \le n-2\}$ and then $\kappa(L) = n$. If L is infinite, then $L = a^*$ and $\kappa(L) = 1$. The bounds for unary languages are in Tables 1 and 2.

5 Kuratowski Algebras Generated by Closed Regular Languages

A theorem of Kuratowski [24] states that, given a topological space, at most 14 distinct sets can be produced by repeatedly applying the operations of closure and complement to a given set. A *closure* operation on a set S is an operation $\square : 2^S \to 2^S$ satisfying the following conditions for any subsets X, Y of S:

$$X \subseteq X^{\square}; \qquad X \subseteq Y \text{ implies } X^{\square} \subseteq Y^{\square}; \qquad X^{\square\square} \subseteq X^{\square}.$$

Kuratowski's theorem was studied in the setting of formal languages in [7]. Positive closure and Kleene closure (star) are both closure operations. It then follows that at most 10 distinct languages can be produced by repeatedly applying the operations of positive closure and complement to a given language, and at most 14 distinct languages can be produced with Kleene closure instead of positive closure. Here we consider the case where the given language is closed and regular, and give upper bounds on the quotient complexity of the resulting languages. In this section, we denote the complement of a language L by L^-, the positive closure of the complement of L by L^{-+}, etc.

We begin with positive closure. Let L be a $\trianglelefteq$-closed language not equal to Σ^*. Then L^- is an ideal, and $L^{-+} = L^-$. In addition, L^+ is also $\trianglelefteq$-closed, so $L^{+-+} = L^{+-}$. Hence there are at most 4 distinct languages that can be produced with positive closure and complementation.

Theorem 6 (Positive Closure) *The worst-case complexities in the* 4*-element algebra generated by a closed language* L *with* $\kappa(L)=n$ *under positive closure and complement are* $\kappa(L)=\kappa(L^-)=n$, $\kappa(L^+)=\kappa(L^{+-})=f(n)$, *where*

1. $f(n)=2^{n-2}+1$ *for prefix-closed languages*,
2. $f(n)=n-1$ *for suffix-closed languages*,
3. $f(n)=2$ *for factor- and subword-closed languages*.

There exist closed languages that meet these bounds.

Proof Since $L^+=L^*$ for a non-empty closed language we have $\kappa(L^+)=\kappa(L^*)$, and the upper bounds $f(n)$ follow from our results on the quotient complexity of the star operation; in the case of suffix-closed languages, to get a 4-element algebra we need $L\neq L^*$. All the languages that we have used in Theorem 4 to prove tightness can be used as examples meeting the bound $f(n)$. □

Kleene closure is similar. Let L be a non-empty $\trianglelefteq$-closed language not equal to Σ^*. Then L^- is an ideal and L^- does not contain ε. Thus $L^{-*}=L^-\cup\varepsilon$ and $L^{-*-}=L\setminus\varepsilon$, which gives at most four languages thus far. Now $L^*=(L\setminus\varepsilon)^*$, and the language L^* is also $\trianglelefteq$-closed. By the previous reasoning, we have at most four additional languages, giving a total of eight languages as the upper bound. The 8-element algebras are of the form $(L, L^-, L^{-*}=L^-\cup\varepsilon, L^{-*-}=L\setminus\varepsilon, L^*, L^{*-}, L^{*-*}=L^{*-}\cup\varepsilon, L^{*-*-}=L^*\setminus\varepsilon)$.

Theorem 7 (Kleene Closure) *The worst-case quotient complexities in the* 8*-element algebra generated by a closed language* L *with* $\kappa(L)=n$ *under star and complement are as follows*: $\kappa(L)=\kappa(L^-)=n$, $\kappa(L^*)=\kappa(L^{*-})=f(n)$, $\kappa(L^{*-*})=\kappa(L^{*-*-})=f(n)+1$, $\kappa(L^{-*})=\kappa(L^{-*-})=n+1$, *where* $f(n)$ *is defined as in Theorem* 6. *There exist closed languages that meet these bounds*.

Proof Since $L^{-*-}=L\setminus\varepsilon$ and $L^{*-*-}=L^*\setminus\varepsilon$ we have $\kappa(L^{-*-})\leq n+1$ and $\kappa(L^{*-*-})\leq f(n)+1$. In the case of suffix-closed languages, since L must be distinct from L^*, we have $f(n)=n-1$ by Theorem 4.

1. Let L be the prefix-closed language defined by the minimal dfa in Fig. 5; then L meets the upper bound on star. Add a loop with a new letter d in each state and denote the resulting language by K. Then K is a prefix-closed language with $\kappa(K)=n$ and $\kappa(K\setminus\varepsilon)=n+1$. Next we have $\kappa(K^*)=\kappa(L^*)=2^{n-2}+1$ and $\kappa(K^*\setminus\varepsilon)=2^{n-2}+2$.

2. Let $L=b^*\cup\bigcup_{i=1}^{n-3}b^*a^ib$. Then L is a suffix-closed language with $\kappa(L)=n$ and $\kappa(L\setminus\varepsilon)=n+1$. Next $\kappa(L^*)=n-1$ and $\kappa(L^*\setminus\varepsilon)=n$.

3. Let $L=\{w\in\{a,b,c\}^*\mid w=b^*a^i \text{ and } 0\leq i\leq n-2\}$. Then L is a subword-closed language with $\kappa(L)=n$ and $\kappa(L\setminus\varepsilon)=n+1$. Next $L^*=\{a,b\}^*$, and so $\kappa(L^*)=2$ and $\kappa(L^*\setminus\varepsilon)=3$. □

Table 1 Bounds on quotient complexity of boolean operations

	$K \cup L$	$\|\Sigma\|$	$K \cap L$	$\|\Sigma\|$	$K \setminus L$	$\|\Sigma\|$	$K \oplus L$	$\|\Sigma\|$
Unary closed	$\max(m,n)$	1	$\min(m,n)$	1	m	1	$\max(m,n)$	1
prefix-, factor-, subword-closed	mn	2	$mn-(m+n-2)$	2	$mn-(n-1)$	2	mn	2
suffix-closed	mn	4	mn	2	mn	4	mn	2
regular	mn	2	mn	2	mn	2	mn	2

Table 2 Bounds on quotient complexity of closure, product, star and reversal

	$\trianglelefteq L$	$\|\Sigma\|$	KL	$\|\Sigma\|$	K^*	$\|\Sigma\|$	K^R	$\|\Sigma\|$
Unary closed	n	1	$m+n-2$	1	2	1	n	1
prefix–closed	n	1	$(m+1)2^{n-2}$	3	$2^{n-2}+1$	3	2^{n-1}	2
suffix-closed	2^n-1	2	$(m-1)n+1$	3	n	2	$2^{n-1}+1$	3
factor-closed	2^{n-1}	2	$m+n-1$	2	2	2	$2^{n-2}+1$	3
subword-closed	$2^{n-2}+1$	$n-2$	$m+n-1$	2	2	2	$2^{n-2}+1$	$2n$
regular	–	–	$m2^n-2^{n-1}$	2	$2^{n-1}+2^{n-2}$	2	2^n	2

6 Conclusions

Tables 1 and 2 summarize our complexity results. The complexities for regular languages are from [22, 25, 26, 32], except those for difference and symmetric difference, which are from [6]. The bounds for boolean operations and reversal of closed languages are direct consequences of the results in [8]. The tables also show the size of the alphabet of the witness languages. In all cases when the size of the alphabet is more than two, we do not know whether the bounds are tight for a smaller alphabet.

References

1. Ang, T., Brzozowski, J.: Languages convex with respect to binary relations, and their closure properties. Acta Cybern. **19**(2), 445–464 (2009)
2. Avgustinovich, S.V., Frid, A.E.: A unique decomposition theorem for factorial languages. Int. J. Algebra Comput. **15**, 149–160 (2005)
3. Bassino, F., Giambruno, L., Nicaud, C.: Complexity of operations on cofinite languages. In: López-Ortiz, A. (ed.) Proceedings of the 9th Latin American Theoretical Informatics Symposium (LATIN). LNCS, vol. 6034, pp. 222–233. Springer, Berlin (2010)
4. Brzozowski, J.: Derivatives of regular expressions. J. ACM **11**(4), 481–494 (1964)
5. Brzozowski, J.: Complexity in convex languages. In: Dediu, A.H., Fernau, H., Martin-Vide, C. (eds.) Proceedings of the 4th International Conference on Language and Automata Theory (LATA). LNCS, vol. 6031, pp. 1–15. Springer, Berlin (2010)
6. Brzozowski, J.: Quotient complexity of regular languages. J. Autom. Lang. Comb. **15**(1/2), 71–89 (2010)
7. Brzozowski, J., Grant, E., Shallit, J.: Closures in formal languages and Kuratowski's theorem. Int. J. Found. Comput. Sci. **22**, 301–321 (2011)

8. Brzozowski, J., Jirásková, G., Li, B.: Quotient complexity of ideal languages. Theor. Comput. Sci. **470**, 36–52 (2013)
9. Brzozowski, J., Jirásková, G., Li, B., Smith, J.: Quotient complexity of bifix-, factor-, and subword-free languages. In: Dömösi, P., Szabolcs, I. (eds.) Proceedings of the 13th Int. Conference on Automata and Formal Languages (AFL), pp. 123–137. Institute of Mathematics and Informatics, College of Nyíregyháza, Nyíregyháza (2011)
10. Brzozowski, J., Santean, N.: Predictable semiautomata. Theor. Comput. Sci. **410**, 3236–3249 (2009)
11. Brzozowski, J., Shallit, J., Xu, Z.: Decision problems for convex languages. Inf. Comput. **209**, 353–367 (2011)
12. Câmpeanu, C., Culik, K. II, Salomaa, K., Yu, S.: State complexity of basic operations on finite languages. In: Boldt, O., Jürgensen, H. (eds.) Revised papers from the 4th International Workshop on Automata Implementation, (WIA). LNCS, vol. 2214, pp. 60–70. Springer, Berlin (2001)
13. 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, Berlin (1990)
14. Galil, Z., Simon, J.: A note on multiple-entry finite automata. J. Comput. Syst. Sci. **12**, 350–351 (1976)
15. Gill, A., Kou, L.T.: Multiple-entry finite automata. J. Comput. Syst. Sci. **9**(1), 1–19 (1974)
16. Gruber, H., Holzer, M., Kutrib, M.: The size of Higman-Haines sets. Theor. Comput. Sci. **387**, 167–176 (2007)
17. Gruber, H., Holzer, M., Kutrib, M.: More on the size of Higman-Haines sets: effective constructions. Fundam. Inform. **91**(1), 105–121 (2009)
18. Haines, L.H.: On free monoids partially ordered by embedding. J. Comb. Theory **6**(1), 94–98 (1969)
19. Han, Y.S., Salomaa, K.: State complexity of basic operations on suffix-free regular languages. Theor. Comput. Sci. **410**(27–29), 2537–2548 (2009)
20. 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, pp. 99–115. University of Szeged, Szeged (2009)
21. Holzer, M., Salomaa, K., Yu, S.: On the state complexity of k-entry deterministic finite automata. J. Autom. Lang. Comb. **6**, 453–466 (2001)
22. Jirásek, J., Jirásková, G., Szabari, A.: State complexity of concatenation and complementation. Int. J. Found. Comput. Sci. **16**, 511–529 (2005)
23. Kao, J.Y., Rampersad, N., Shallit, J.: On NFAs where all states are final, initial, or both. Theor. Comput. Sci. **410**(47–49), 5010–5021 (2009)
24. Kuratowski, C.: Sur l'opération $\overline{A}$ de l'analysis situs. Fundam. Math. **3**, 182–199 (1922) (in French)
25. Maslov, A.N.: Estimates of the number of states of finite automata. Dokl. Akad. Nauk SSSR **194**, 1266–1268 (1970) (in Russian). English translation: Soviet Math. Dokl. **11**, 1373–1375 (1970)
26. Mirkin, B.G.: On dual automata. Kibernetika **2**, 7–10 (1966) (in Russian). English translation: Cybernetics **2**, 6–9 (1970)
27. Okhotin, A.: On the state complexity of scattered substrings and superstrings. Fundam. Inform. **99**(3), 325–338 (2010)
28. Pighizzini, G., Shallit, J.: Unary language operations, state complexity and Jacobsthal's function. Int. J. Found. Comput. Sci. **13**, 145–159 (2002)
29. Thierrin, G.: Convex languages. In: Nivat, M. (ed.) Automata, Languages and Programming, pp. 481–492. North-Holland, Amsterdam (1973)
30. Veloso, P.A.S., Gill, A.: Some remarks on multiple-entry finite automata. J. Comput. Syst. Sci. **18**, 304–306 (1979)
31. Yu, S.: State complexity of regular languages. J. Autom. Lang. Comb. **6**, 221–234 (2001)
32. Yu, S., Zhuang, Q., Salomaa, K.: The state complexities of some basic operations on regular languages. Theor. Comput. Sci. **125**(2), 315–328 (1994)