Upper Bounds on Syntactic Complexity of Left and Two-Sided Ideals

Marek Szykuła

University of Wrocław, Poland

Joint work with Janusz Brzozowski (University of Waterloo, Canada)

DLT, 27.08.2014

Abstract

We study syntactic complexity of two subclasses of regular languages:

- left ideals $L = \Sigma^* L$,
- two-sided ideals $L = \Sigma^* L \Sigma^*$.

Contribution

- Brzozowski and Ye (DLT 2011) conjectured that:
  - The syntactic complexity of left ideals (and suffix-closed languages) is $n^{n-1} + n - 1$.
  - The syntactic complexity of two-sided ideals (and factor-closed languages) is $n^{n-2} + (n-2)2^{n-2} + 1$ (for $n > 1$).
- We prove these conjectures.

Left quotient

The (left) quotient of a regular language L by a word w is

$$Lw = \{x \in \Sigma^* \mid wx \in L\}.$$

Analogously, for a state q of a minimal DFA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ recognizing L:

$$L_q = \{x \in \Sigma^* \mid qt_x \in F\},$$

where t_x is the transformation of word x.

So L_q is the set of words taking q to an accepting state.

Left quotient

The (left) quotient of a regular language L by a word w is

$$Lw = \{x \in \Sigma^* \mid wx \in L\}.$$

Analogously, for a state q of a minimal DFA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ recognizing L:

$$L_q = \{x \in \Sigma^* \mid qt_x \in F\},$$

where t_x is the transformation of word x.

So L_q is the set of words taking q to an accepting state.

Left quotient

The (left) quotient of a regular language L by a word w is

$$Lw = \{x \in \Sigma^* \mid wx \in L\}.$$

Analogously, for a state q of a minimal DFA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ recognizing L:

$$L_q = \{x \in \Sigma^* \mid qt_x \in F\},$$

where t_x is the transformation of word x.

So L_q is the set of words taking q to an accepting state.

State complexity

Nerode right congruence on Σ^*

For a regular language L and words $x, y \in \Sigma^*$:

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

State complexity

The state complexity or quotient complexity $\kappa(L)$ of a regular language L is:

- The number of equivalence classes of $\sim_L$.
- The number of left quotients of L.
- The number of states in a minimal DFA recognizing L.

Syntactic complexity

Myhill congruence

For a regular language L and words $x, y \in \Sigma^*$:

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

Syntactic complexity

The syntactic complexity $\sigma(L)$ of a regular language L is:

- $|\Sigma^+/\approx_L|$ – the number of equivalence classes of $\approx_L$.
- The size of the syntactic semigroup of L.
- The size of the transition semigroup of a minimal DFA recognizing L.

Syntactic complexity of a class of languages

The syntactic complexity of a class of languages is:

- The size of the largest syntactic semigroups of languages in that class.
- Expressed as a function of the state complexities $n = \kappa(L)$ of the languages.

In other words

- Suppose we have an n-state minimal DFA recognizing some language from the given class.
- We ask how many transformations (at most) can be in the transition semigroup of the DFA.

Syntactic complexity of a class of languages

The syntactic complexity of a class of languages is:

- The size of the largest syntactic semigroups of languages in that class.
- Expressed as a function of the state complexities $n = \kappa(L)$ of the languages.

In other words

- Suppose we have an n-state minimal DFA recognizing some language from the given class.
- We ask how many transformations (at most) can be in the transition semigroup of the DFA.

Proposition

- $n-1 \leq \sigma(L) \leq n^n$
- *The bounds are tight for $n > 1$ in the class of all regular languages.*

Previous results

- Gomes, Howie 1992: (partially) monotonic semigroups.
- Krawetz, Lawrence, Shallit 2003: unary and binary alphabets.
- Holzer, König 2004: unary and binary alphabets.
- Brzozowski, Ye 2010: ideal and closed languages.
- Beaudry, Holzer 2011: semigroups of reversible DFAs.
- Brzozowski, Liu 2012: finite, cofinite, definite, reverse definite languages.
- Brzozowski, Li, Ye 2012: prefix-, suffix-, bifix-, factor-free languages.
- Iván, Nagy-György 2013: (generalized) definite languages.
- Brzozowski, Li 2013: $\mathcal{J}$-trivial and $\mathcal{R}$-trivial languages.
- Brzozowski, Li, Liu 2013: aperiodic, nearly monotonic semigroups.
- Brzozowski, Szykuła 2014: aperiodic.

Ideals

Ideals

- Right ideal $L = L\Sigma^*$.
- Left ideal $L = \Sigma^* L$.
- Two-sided ideal $L = \Sigma^* L \Sigma^*$.

Closed languages

- Right ideals are complements of prefix-closed languages.
- Left ideals are complements of suffix-closed languages.
- Two-sided ideals are complements of factor-closed languages.

Syntactic complexity is preserved under complementation, so our proofs are in terms of ideals only.

Theorem (Brzozowski, Ye 2010)

- *Syntactic complexity of right ideals is equal to* n^{n-1}
- *Syntactic complexity of left ideals is at least* $n^{n-1} + n - 1$
- *Syntactic complexity of two-sided ideals is at least* $n^{n-2} + (n-2)2^{n-2} + 1$ *(for* $n \geq 2$*)*

We have proved that the lower bounds for syntactic complexity of left ideals and two-sided ideals are also upper bounds.

Left ideals

Order of quotients

We define a partial order $\preceq$ on the set of states Q:

$$p \preceq q \text{ if and only if } L_p \subseteq L_q.$$

In other words:

- If a transformation maps p to an accepting state, then it must also map q to an accepting one.
- $p \prec q$ means that q is "closer" to an accepting state than p.

Proposition

If $p \preceq q$, then $pt \preceq qt$ for any transformation t.

Order of quotients

We define a partial order $\preceq$ on the set of states Q:

$$p \preceq q \text{ if and only if } L_p \subseteq L_q.$$

In other words:

- If a transformation maps p to an accepting state, then it must also map q to an accepting one.
- $p \prec q$ means that q is "closer" to an accepting state than p.

Proposition

If $p \preceq q$, then $pt \preceq qt$ for any transformation t.

Properties of $\preceq$

Proposition (Left Ideals)

Initial state $0 \preceq q$ *for any state* q.

This is the characterization of minimal DFAs recognizing left ideals.

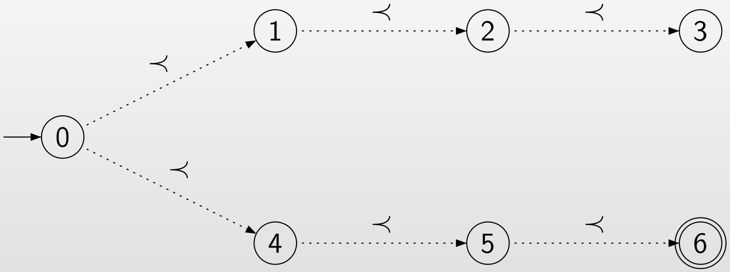


Figure : Example partial order $\preceq$.

Here $0 \prec 1 \prec 2 \prec 3$ and $0 \prec 4 \prec 5 \prec 6$.

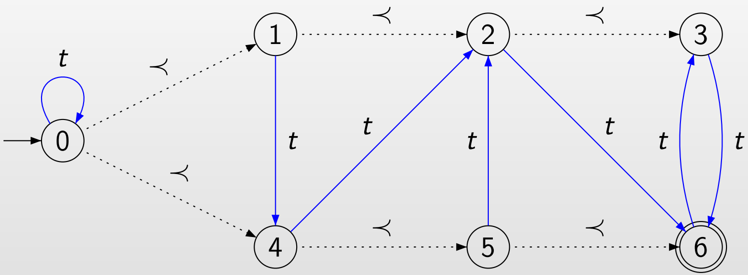


Figure : Allowed transformation t preserving $\preceq$.

We have $p \preceq q$ implies $pt \preceq qt$.

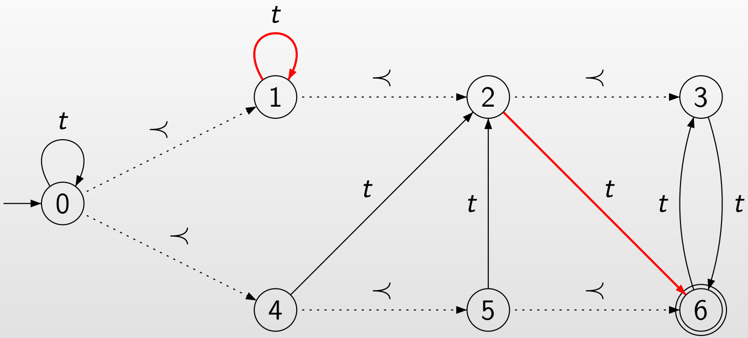


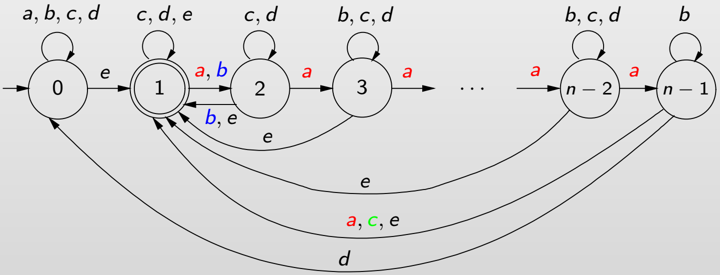
Figure : Disallowed transformation t violating the ordering.

Here $1 \prec 2$, but $1t = 1 \not\prec 6 = 2t$.

In particular

- Every cycle in a transformation consists of only pairwise incomparable states under $\preceq$.
- Every maximal chain $p \prec pt \prec pt^2 \prec \ldots \prec pt^k$ with $k \geq 1$ ends with a fixed point ($pt^k = pt^{k+1}$).

(Brzozowski, Ye 2010) Witness DFA with $n^{n-1} + n - 1$ transformations



The transition semigroup S_n of the witness contains:

- All transformations that fix 0; the other states are mapped arbitrarily.
- All constant transformations.

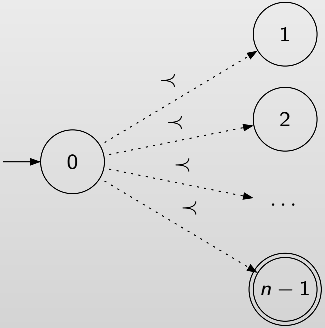


Figure : The partial order $\prec$ in S_n.

Upper bound

- S_n – the transition semigroup of the witness.
- T_n – the transition semigroup of an arbitrary left ideal.
- We show that $|T_n| \leq |S_n| = n^{n-1} + n - 1$.

Idea

- It is possible that $T_n \not\subseteq S_n$.
- We do not count $|T_n|$ directly.
- We construct an injective function of transformations

$$f : T_n \to S_n$$

- (Injective: for every $t \neq t'$ we have $f(t) \neq f(t')$.)

Upper bound

- S_n – the transition semigroup of the witness.
- T_n – the transition semigroup of an arbitrary left ideal.
- We show that $|T_n| \leq |S_n| = n^{n-1} + n - 1$.

Idea

- It is possible that $T_n \not\subseteq S_n$.
- We do not count $|T_n|$ directly.
- We construct an injective function of transformations

$$f\colon T_n \to S_n$$

- (Injective: for every $t \neq t'$ we have $f(t) \neq f(t')$.)

Upper bound

- S_n – the transition semigroup of the witness.
- T_n – the transition semigroup of an arbitrary left ideal.
- We show that $|T_n| \leq |S_n| = n^{n-1} + n - 1$.

Idea

- It is possible that $T_n \not\subseteq S_n$.
- We do not count $|T_n|$ directly.
- We construct an injective function of transformations

$$f\colon T_n \to S_n$$

- (Injective: for every $t \neq t'$ we have $f(t) \neq f(t')$.)

Upper bound

- S_n – the transition semigroup of the witness.
- T_n – the transition semigroup of an arbitrary left ideal.
- We show that $|T_n| \le |S_n| = n^{n-1} + n - 1$.

Idea

- It is possible that $T_n \not\subseteq S_n$.
- We do not count $|T_n|$ directly.
- We construct an injective function of transformations

$$f\colon T_n \to S_n$$

- (Injective: for every $t \neq t'$ we have $f(t) \neq f(t')$.)

Case 1

If $t \in S_n$, then let $f(t) = t$.

This is obviously injective, and $f(t) \in T_n$.

From now, for $t \notin S_n$ we need to assign a transformation $f(t) \notin T_n$.

Case 1

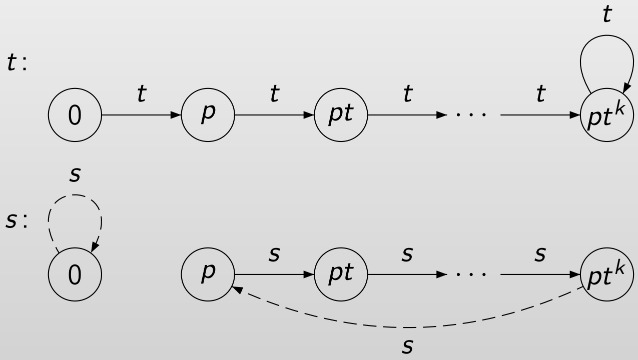
If $t \in S_n$, then let $f(t) = t$.

This is obviously injective, and $f(t) \in T_n$.

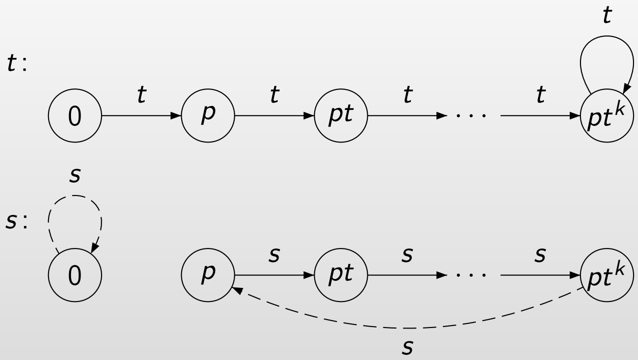
From now, for $t \notin S_n$ we need to assign a transformation $f(t) \notin T_n$.

Case 2

If $t \notin S_n$ and $0t \neq 0t^2$, then we define $f(t) = s$ as follows:



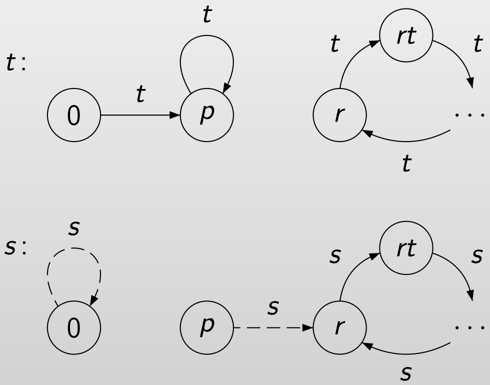
Case 2



- $0 \prec p \prec pt \prec \ldots \prec pt^k$, so s violates $\prec$ in T_n.
- $s \notin T_n$ (in opposite to Case 1), but $s \in S_n$ (since 0 is fixed).
- $f(t)$ is injective in this case (given s, we may reproduce t unambiguously).

Case 3(a)

If $t \notin S_n$ and $0t^2 = 0t$, and t has a cycle, then:



Summary

We have 5 (sub)cases

Case 1 $t \in S_n$.

Case 2 $t \notin S_n$ and $0t^2 \neq 0t$.

Case 3 $t \notin S_n$ and $0t^2 = 0t$.

(a) t has a cycle.

(b) t has no cycles, but has a fixed point $\neq 0t$.

(c) t has no cycles nor fixed points $\neq 0t$, but there is a state $r \succ 0t$ mapped to $0t$.

These cover all possibilities for t.
So f is injective and $f(T_n) \subseteq S_n$, and $|T_n| = |f(T_n)| \leq |S_n|$.

Uniqueness of maximality

The transition semigroup S_n of the witness is the only one reaching the upper bound:

Theorem

If $n \geq 3$, $\mathcal{A} = (\{0, \dots, n-1\}, \Sigma, 0, F, \delta)$ is a minimal DFA of a left ideal, and its transition semigroup T_n has size $n^{n-1} + n - 1$, then

$$T_n = S_n.$$

Two-sided ideals

Two-sided ideals

- A two-sided ideal is simultaneously left and right ideal.

So in addition to the properties of DFAs of left ideals we have:

The properties of right ideals

In minimal DFAs recognizing right ideals:

- There is only one final state, say $n-1$.
- Every transformation t fixes $n-1$ (state $n-1$ is a *sink*).
- $0 \preceq q \preceq n-1$ for every state q.

Two-sided ideals

- A two-sided ideal is simultaneously left and right ideal.

So in addition to the properties of DFAs of left ideals we have:

The properties of right ideals

In minimal DFAs recognizing right ideals:

- There is only one final state, say $n-1$.
- Every transformation t fixes $n-1$ (state $n-1$ is a *sink*).
- $0 \preceq q \preceq n-1$ for every state q.

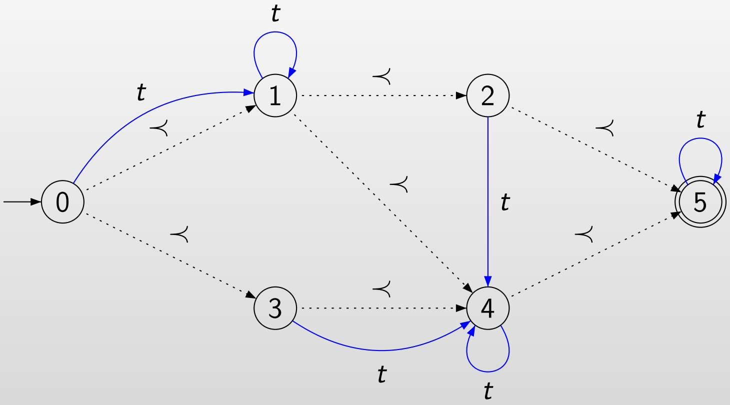
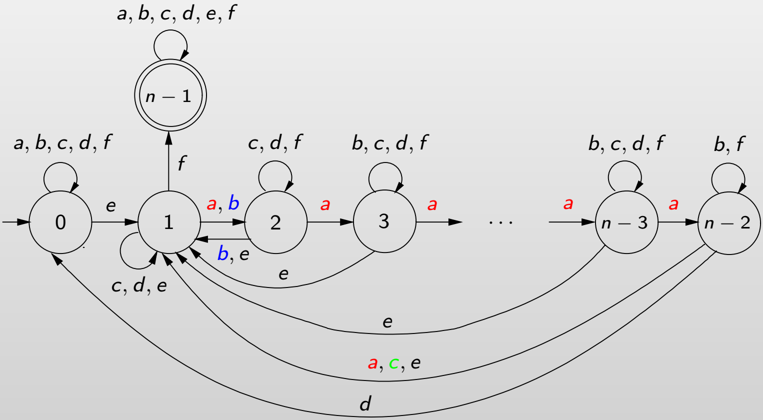


Figure : Allowed transformation t preserving $\preceq$.

(Brzozowski, Ye 2010)
Witness DFA with $n^{n-2} + (n-2)2^{n-2} + 1$ transformations



The transition semigroup S_n of the witness contains:

- All transformations from $\{1, \dots, n-2\}$ to $\{1, \dots, n-1\}$ fixing 0 and $n-1$.
- All transformations mapping $\{0, \dots, n-2\}$ to a single state and fixing $n-1$.

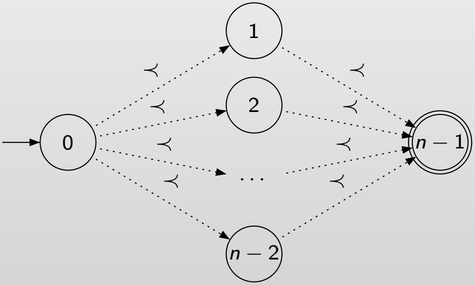


Figure : The partial order $\prec$ in S_n.

Upper bound

Again, the proof follows by constructing an injective function

$$f\colon\; T_n \to S_n.$$

Additional difficulty

- We may reuse some of the cases from the proof for left ideals.
- But now, final state $n-1$ must be fixed by all defined transformations.

Upper bound

Again, the proof follows by constructing an injective function

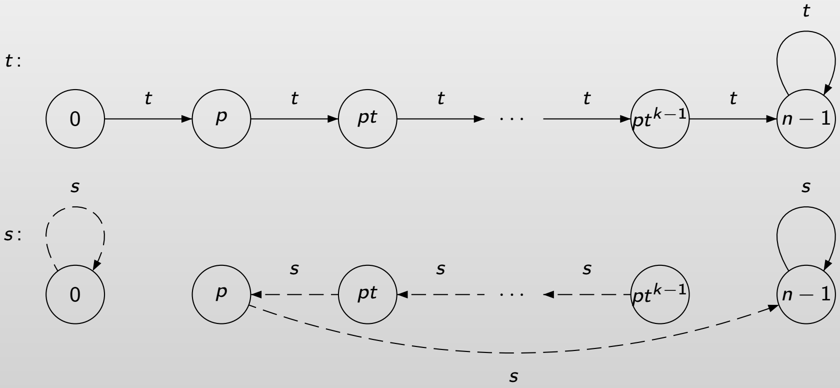
$$f\colon T_n \to S_n.$$

Additional difficulty

- We may reuse some of the cases from the proof for left ideals.
- But now, final state $n-1$ must be fixed by all defined transformations.

Case 2(b)

If $t \notin S_n$ and $0t^2 \neq 0t$, then we define $f(t) = s$ as follows:



Summary

We have 8 (sub)cases

Case 1 $t \in S_n$.

Case 2 $t \notin S_n$ and $0t^2 \neq 0t$.

- (a) Fixed point $0t^k \neq n-1$.
- (b) Fixed point $0t^k = n-1$ and $k \geq 3$.
- (c) $0t = n-1$.

Case 3 $t \notin S_n$ and $0t^2 = 0t$.

- (a) t has a cycle.
- (b) t has no cycles, but has a fixed point $\notin \{0t, n-1\}$.
- (c) t has no cycles nor fixed points $\notin \{0t, n-1\}$, but there is a state $r \succ 0t$ mapped to $0t$.
- (d) t has no cycles, no fixed points $\notin \{0t, n-1\}$, and no states $r \succ 0t$ mapped to $0t$, but there is a state $r \succ 0t$ mapped to $n-1$.

Uniqueness of maximality

Again, the transition semigroup S_n of the witness is the only one reaching the upper bound:

Theorem

If $n \geq 4$, $\mathcal{A} = (\{0, \ldots, n-1\}, \Sigma, 0, F, \delta)$ *is a minimal DFA of a two-sided ideal, and its transition semigroup* T_n *has size* $n^{n-2} + (n-2)2^{n-2} + 1$, *then*

$$T_n = S_n.$$

Future work

Other problems

The technique of injective functions may be applied to solve similar problems.
For example, the upper bounds for syntactic complexity of suffix-, bifix-, and factor-free languages.

Suffix-free

- We have already the proof of the tight upper bound for suffix-free languages! (12 cases)

Большое Спасибо!

Future work

Other problems

The technique of injective functions may be applied to solve similar problems.
For example, the upper bounds for syntactic complexity of suffix-, bifix-, and factor-free languages.

Suffix-free

- We have already the proof of the tight upper bound for suffix-free languages! (12 cases)

Большое Спасибо!

Future work

Other problems

The technique of injective functions may be applied to solve similar problems.
For example, the upper bounds for syntactic complexity of suffix-, bifix-, and factor-free languages.

Suffix-free

- We have already the proof of the tight upper bound for suffix-free languages! (12 cases)

Большое Спасибо!