Descriptional Complexity of Power and Positive Closure on Convex Languages

Michal Hospodár

Mathematical Institute
Slovak Academy of Sciences
Košice, Slovakia

CIAA 2019

July 25, 2019, Košice, Slovakia

Basic Notions

Regular Operations

- Concatenation:
  $KL = \{uv \mid u \in K,\ v \in L\}$
- k-th power:
  $L^k = LL^{k-1}$
  where $L^0 = \{\varepsilon\}$
- Kleene closure:
  $L^* = \bigcup_{i \geq 0} L^i$
- Positive closure:
  $L^+ = \bigcup_{i \geq 1} L^i$

State Complexity

- of a language L, $\mathrm{sc}(L)$,
  is the number of states
  in the minimal DFA for L
- of a unary operation $\circ$:
  $n \mapsto \max\{\mathrm{sc}(L^\circ) \mid \mathrm{sc}(L) \leq n\}$
- of a unary operation $\circ$ on a class $\mathcal{C}$:

$$n \mapsto \max\{\mathrm{sc}(L^\circ) \mid \mathrm{sc}(L) \leq n \text{ and } L \in \mathcal{C}\}$$

Nondeterministic State Complexity

$\mathrm{nsc}(L)$: defined analogously using NFA with a unique initial state

Subclasses of Convex Languages

Prefix, Suffix, Factor, Subword

$w = uxv$

- u is a prefix of w
- v is a suffix of w
- x is a factor of w

$w = u_0 v_1 u_1 \cdots v_m u_m$

- $v_1 v_2 \cdots v_m$ is a subword of w

Free, Closed, Convex

- L is prefix-free if $w \in L$ $\Rightarrow$ no proper prefix of w is in L
- L is prefix-closed if $w \in L$ $\Rightarrow$ every prefix of w is in L
- L is prefix-convex if $u, v \in L$ and u is a prefix of v $\Rightarrow$ every w satisfying
  - u is a prefix of w
  - w is a prefix of v

  is in L

suffix-, factor-, subword- analogously

Subclasses of Convex Languages

Ideal languages

- L is a right ideal if $L = L\Sigma^*$
- L is a left ideal if $L = \Sigma^* L$
- L is a two-sided ideal if $L = \Sigma^* L \Sigma^*$
- L is an all-sided ideal if $L = L \sqcup\!\sqcup \Sigma^*$

Inclusions

$3 \subsetneq 1$
$3 \subsetneq 2$
$4 \subsetneq 3$
$A \subsetneq D$
$B \subsetneq D$
$C \subsetneq D$

	A	B	C	D
1	right ideal	prefix-free	prefix-closed	prefix-convex
2	left ideal	suffix-free	suffix-closed	suffix-convex
3	two-sided ideal	factor-free	factor-closed	factor-convex
4	all-sided ideal	subword-free	subword-closed	subword-convex

Known Results on (Deterministic) State Complexity I.

- Han, Salomaa, Wood: state complexity of basic operations (intersection, union, concatenation, Kleene star) on
  - prefix-free languages (2006)
  - suffix-free languages (2009)
- Jirásková et al.: improvement of sc results on
  - prefix-free languages (2010 with Krausová)
  - suffix-free languages (2009 with Olejár, 2011 with Cmorik)
  - left and right quotient and reversal on prefix-free languages (2016 with Jirásek, Krausová, Mlynárčik, and Šebej)

class \ sc	$K \cap L$	$\lvert\Sigma\rvert$	$K \cup L$	$\lvert\Sigma\rvert$
prefix-free	$mn - 2m - 2n + 6$	2	$mn - m - n + 2$	2
suffix-free	$mn - 2m - 2n + 6$	2	$mn - 2$	2

class \ sc	KL	$\lvert\Sigma\rvert$	L^*	$\lvert\Sigma\rvert$	L^R	$\lvert\Sigma\rvert$
prefix-free	$m + n - 2$	1	m	2	$2^{m-2} + 1$	3
suffix-free	$(m-1)2^{n-2} + 1$	2	$2^{m-2} + 1$	2	$2^{m-2} + 1$	3

Known Results on (Deterministic) State Complexity II.

- Brzozowski, Jirásková, Li, Smith, Zou:
  - Complexity in Convex Languages (2010)
  - Quotient Complexity of Ideal Languages (2013)
  - Quotient Complexity of Closed Languages (2014)
  - Quotient Complexity of Bifix-, Factor-, and Subword-Free Regular Languages (2014)

Quotient complexity = state complexity

Known Results on (Deterministic) State Complexity

class \ sc	$K \cap L$	$K \cup L$	KL	L^*	L^R
right ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
left ideal	✓ ✓	✓, 4	✓ ✓	✓ ✓	✓, 3
two-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3
all-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2n
prefix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
suffix-free	✓ ✓	✓ ✓	✓ ✓	✓, 3	✓ ✓
factor-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3
subword-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
prefix-closed	✓ ✓	✓ ✓	✓, 3	✓, 3	✓ ✓
suffix-closed	✓ ✓	✓, 4	✓, 3	✓ ✓	✓, 3
factor-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3
subword-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2n
prefix-convex	✓ ✓	✓ ✓			
suffix-convex	✓ ✓	✓ ✓			
factor-convex	✓ ✓	✓ ✓			
subword-convex	✓ ✓	✓ ✓			

The first checkmark – the complexity is known

The second checkmark – the complexity is known over an optimal alphabet

The Aims of This Paper I.

class \ sc	$K \cap L$	$K \cup L$	KL	L^*	L^R	L^k	L^+
right ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
left ideal	✓ ✓	✓, 4	✓ ✓	✓ ✓	✓, 3		
two-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3		
all-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2n		
prefix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
suffix-free	✓ ✓	✓ ✓	✓ ✓	✓, 3	✓ ✓		
factor-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3		
subword-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
prefix-closed	✓ ✓	✓ ✓	✓, 3	✓, 3	✓ ✓		
suffix-closed	✓ ✓	✓, 4	✓, 3	✓ ✓	✓, 3		
factor-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3		
subword-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2n		
prefix-convex	✓ ✓	✓ ✓					
suffix-convex	✓ ✓	✓ ✓					
factor-convex	✓ ✓	✓ ✓					
subword-convex	✓ ✓	✓ ✓					

Nondeterministic State Complexity of Operations

History

- Holzer, Kutrib (2003):
  definition of nsc, basic operations on regular languages
- Prefix-free languages: Han, Salomaa, Wood (2009), Jirásková, Krausová (2010)
- Suffix-free: Han, Salomaa (2010), Jirásková, Olejár (2009)
- Mlynárčik et al.:
  complement on prefix-free, suffix-free, non-returning (2014)
  complement on free and ideal (2015)
  basic operations on closed and ideal (2016)
  basic operations on free and convex (2017)

Motivation – the operations of the k-th power and positive closure were not considered in most of these papers

Known Results on Nondeterministic State Complexity

class \ nsc	$K \cap L$	$K \cup L$	KL	L^*	L^R	L^c
right ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
left ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
two-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
all-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2^n
prefix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
suffix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
factor-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
subword-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2^n
prefix-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓
suffix-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓
factor-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓
subword-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓, $2n$	✓, 2^n
prefix-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓
suffix-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓, 5
factor-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	
subword-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓, $2n$	

The Aims of This Paper II.

class \ nsc	$K \cap L$	$K \cup L$	KL	L^*	L^R	L^c	L^k	L^+
right ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
left ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
two-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
all-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2^n		
prefix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
suffix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
factor-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		
subword-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2^n		
prefix-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓		
suffix-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓		
factor-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓		
subword-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓, $2n$	✓, 2^n		
prefix-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓ ✓		
suffix-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓	✓, 5		
factor-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓			
subword-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓, $2n$			

Known Results for L^k and L^+ on Regular Languages

- Rampersad: The state complexity of L^2 and L^k (2006)
- Domaratzki, Okhotin: State complexity of power (2009)
- Holzer, Kutrib: Nondeterministic descriptional complexity of regular languages (2003)

Known results

(Deterministic) state complexity

	L^k	$\lvert\Sigma\rvert$	L^+	$\lvert\Sigma\rvert$
regular	$\Theta(n2^{(k-1)n})$	6	$(3/4)2^n - 1$	2
unary regular	$k(n-1)+1$		$(n-1)^2$	

Nondeterministic state complexity

	L^k	$\lvert\Sigma\rvert$	L^+	$\lvert\Sigma\rvert$
regular	kn	2	n	1
unary regular	$k(n-1)+1 \leq \cdot \leq kn$		n	

A Useful Lemma Used In Our Proof

Definition

$$A = (Q, \Sigma, \cdot, I, F)$$

$$A^R = (Q, \Sigma, \cdot^R, F, I)$$

$S \subseteq Q$ is reachable in A if $S = I \cdot w$ for some w in $L(A)$
$S \subseteq Q$ is co-reachable in A if S is reachable in A^R

Lemma (Greater-Smaller Lemma)

Let A be an NFA with the state set $\{1, 2, \ldots, n\}$.
Let $\{(X_i, Y_i) \mid i = 1, 2, \ldots, n\}$ be a set of pairs s.t. for each i:

1. $X_i \subseteq \{i, i+1, \ldots, n\}$ *and* $Y_i \subseteq \{1, 2, \ldots, i\}$,
2. $i \in X_i \cap Y_i$, *and*
3. X_i *is reachable and* Y_i *is co-reachable in* A.

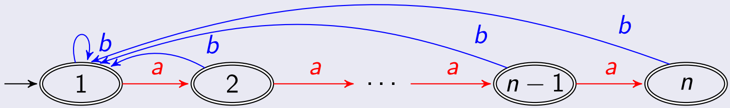
Then A is a minimal NFA.

- this allows us to avoid describing fooling sets

The Most Interesting Result of This Paper I.

Theorem

There exists a binary factor-closed language L accepted by an n-state NFA such that every NFA for L^k has at least kn states.



$L = \{w \in \{a,b\}^* \mid \text{each factor of } w \text{ in } a^* \text{ is of length} \leq n-1\}$

$\Rightarrow L$ is factor-closed

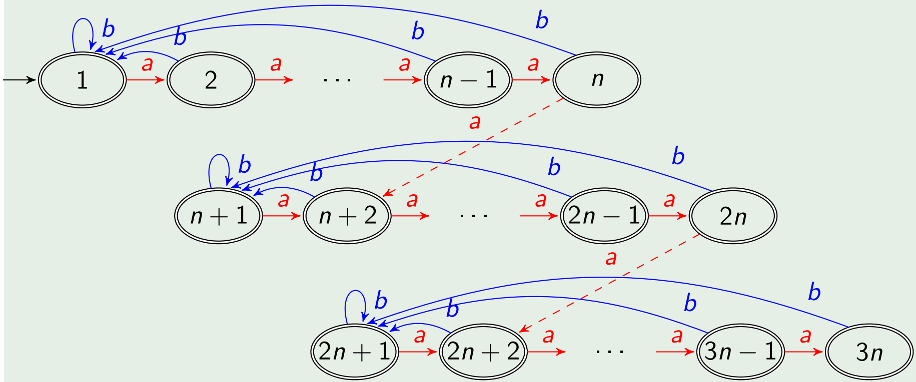
Proof idea: lower bound kn for the k-th power

- the partial DFA D for L^k has kn states
- D satisfies the conditions in lemma from the previous slide
- $\Rightarrow D$ is a minimal NFA for L^k □

The Most Interesting Result of This Paper II.

Example ($k = 3$)

The minimal partial DFA D for L^3



- three copies of A
- connected through the transitions $(n, a, n+2)$ and $(2n, a, 2n+2)$

NSC of L^k and L^+ on Subclasses of Convex Languages

	L^k	$\|\Sigma\|$	L^+	$\|\Sigma\|$
right ideal	$k(n-1)+1$,	1	n,	1
left ideal	$k(n-1)+1$,	1	n,	1
two-sided ideal	$k(n-1)+1$,	1	n,	1
all-sided ideal	$k(n-1)+1$,	1	n,	1
prefix-free	$k(n-1)+1$,	1	n,	1
suffix-free	$k(n-1)+1$,	1	n,	1
factor-free	$k(n-1)+1$,	1	n,	1
subword-free	$k(n-1)+1$,	1	n,	1
prefix-closed	kn,	2	n,	2
suffix-closed	kn,	2	n,	2
factor-closed	kn,	2	1,	1
subword-closed	kn,	3	1,	1
prefix-convex	kn,	2	n,	1
suffix-convex	kn,	2	n,	1
factor-convex	kn,	2	n,	1
subword-convex	kn,	3	n,	1

Summary – Nondeterministic State Complexity

	Known results						Here	
	$K \cap L$	$K \cup L$	KL	L^*	L^R	L^c	L^k	L^+
right ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
left ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
two-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
all-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2^n	✓ ✓	✓ ✓
prefix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
suffix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
factor-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
subword-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2^n	✓ ✓	✓ ✓
prefix-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
suffix-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
factor-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
subword-closed	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓, $2n$	✓, 2^n	✓, 3	✓ ✓
prefix-convex	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
suffix-convex	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 5	✓ ✓	✓ ✓
factor-convex	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓		✓ ✓	✓ ✓
subword-convex	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓, $2n$		✓, 3	✓ ✓

Summary – (Deterministic) State Complexity

	Known results					Here	
class \ sc	$K \cap L$	$K \cup L$	KL	L^*	L^R	L^k	L^+
right ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
left ideal	✓ ✓	✓, 4	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓
two-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓
all-sided ideal	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2n	✓ ✓	✓ ✓
prefix-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
suffix-free	✓ ✓	✓ ✓	✓ ✓	✓, 3	✓ ✓	? ?	✓ ✓
factor-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓
subword-free	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
prefix-closed	✓ ✓	✓ ✓	✓, 3	✓, 3	✓ ✓	? ?	✓ ✓
suffix-closed	✓ ✓	✓, 4	✓, 3	✓ ✓	✓, 3	? ?	✓ ✓
factor-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 3	✓ ✓	✓ ✓
subword-closed	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓, 2n	✓ ✓	✓ ✓
prefix-convex	✓ ✓	✓ ✓					
suffix-convex	✓ ✓	✓ ✓					
factor-convex	✓ ✓	✓ ✓					
subword-convex	✓ ✓	✓ ✓					

Summary and Open Problems

Summary

we get the tight upper bounds for the (nondeterministic) state complexity of power and positive closure

- nsc: everywhere (both operations for prefix-, suffix-, factor-, and subword-free, -closed, and -convex, and right, left, two-sided, and all-sided ideal languages)
- sc: everywhere but for
  - power on suffix-free, prefix-closed, and suffix-closed languages,
  - and proper convex languages

the alphabet is optimal everywhere

- for nondeterministic complexity of power on subword-closed and subword-convex languages, the binary witness was discovered after submitting this paper

Summary and Open Problems

Open Problems

we get the tight upper bounds for the (nondeterministic) state complexity of power and positive closure

- nsc: everywhere (both operations for prefix-, suffix-, factor-, and subword-free, -closed, and -convex, and right, left, two-sided, and all-sided ideal languages)
- sc: everywhere but for
  - power on suffix-free, prefix-closed, and suffix-closed languages,
  - and proper convex languages

the alphabet is optimal everywhere

- for nondeterministic complexity of power on subword-closed and subword-convex languages, the binary witness was discovered after submitting this paper

Thank You For Your Attention

ĎAKUJEM ZA POZORNOSŤ

Köszönöm	Ευχαριστώ	شكرا لكم
תודה	Tack	Hvala
Děkuji	ありがとう	Obrigado
謝謝	Danke	Teşekkür
Grazie	Спасибо	देव बोरम कोरुम