Nondeterministic Complexity of L^k and L^+ on Subclasses of Convex Languages

Michal Hospodár

Mathematical Institute
Slovak Academy of Sciences
Košice, Slovakia

Joint work with Matúš Palmovský

DLT Satellite Workshop
September 6, 2018, Kyoto, Japan

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:
  $L^+ = \bigcup_{i \geq 1} L^i$

Nondeterministic State Complexity

- of a language L, $\mathrm{nsc}(L)$, is the number of states in a minimal NFA for L
- of a unary operation $\circ$:

$$n \mapsto \max\{\mathrm{nsc}(L^\circ) \mid \mathrm{nsc}(L) \leq n\}$$

- of a unary operation $\circ$ on a class $\mathcal{C}$:

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

Subclasses of Convex Languages

Prefix, Suffix, Factor, Subword

$w = uxv$

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

Ideal

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

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 $u, w \in L$ and $u \leq_p w$ $\Rightarrow$ v with $u \leq_p v \leq_p w$ is in L

suffix, factor, subword analogously

- every prefix-free, -closed, and right ideal language is also prefix-convex
- suffix (left), factor (two-sided), subword (all-sided) analogously

Known Results on (Deterministic) State Complexity

- Han et al.:
  - State Complexity of Prefix-Free Regular Languages (2006)
  - State Complexity of Basic Operations on Suffix-Free Regular Languages (TCS 2009)
- Jirásková et al.:
  - State Complexity of Intersection and Union of Suffix-Free Languages and Descriptional Complexity (NCMA 2009)
  - Complexity in Prefix-Free Regular Languages (DCFS 2010)
  - Basic Operations on Binary Suffix-Free Languages (2011)
  - Prefix-free languages: Left and right quotient and reversal (TCS 2016)
- Brzozowski et al.:
  - Complexity in Convex Languages (LATA 2010)
  - Quotient Complexity of Ideal Languages (TCS 2013)
  - Quotient Complexity of Closed Languages (ToCS 2014)
  - Quotient Complexity of Bifix-, Factor-, and Subword-Free Regular Languages (Acta Cybernetica 2014)

Known Results on (Deterministic) State Complexity

	$K \cap L$	$K \cup L$	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	✓ ✓	✓ ✓			

Motivation and History

- Holzer, Kutrib (IJFCS 2003):
  definition of NSC, basic operations on regular languages
- Han, Salomaa, Wood (FI 2009): prefix-free
- Han, Salomaa (DCFS 2010): suffix-free
- Jirásková, Krausová (DCFS 2010): prefix-free
- Jirásková, Olejár (NCMA 2009): boolean op. on suffix-free
- Jirásková, Mlynárčik (DCFS 2014):
  complement on prefix-free, suffix-free, non-returning
- Mlynárčik (DCFS 2015): complement on free and ideal
- Hospodár, Jirásková, Mlynárčik (CIAA 2016): closed, ideal
- Hospodár, Jirásková, Mlynárčik (CIAA 2017): free, convex

Known Results on NSC on Subclasses of Convex Languages

	$K \cap L$	$K \cup L$	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

	$K \cap L$	$K \cup L$	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^+

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

Known results

(Deterministic) state complexity

	L^k	L^+
regular	$\Theta(n2^{(k-1)n})$, $\|\Sigma\| \geq 6$	$\frac{3}{4}2^n - 1$
unary regular	$k(n-1)+1$	$(n-1)^2$

Nondeterministic state complexity

	L^k	L^+
regular	kn, $\|\Sigma\| \geq 2$	n
unary regular	$k(n-1)+1 \leq \cdot \leq kn$	n

A Useful Lemma Used In Our Proof

Lemma 3.

Let $\{(X_i, Y_i) \mid i = 1, 2, \ldots, m\}$ be a set of pairs of subsets of the state set of an NFA A such that for each i in $\{1, 2, \ldots, m\}$

(1) X_i is reachable and Y_i is co-reachable in A,

(2) $i \in X_i \cap Y_i$, and

(3) $X_i \subseteq \{i, i+1, \ldots, n\}$ and $Y_i \subseteq \{1, 2, \ldots, i\}$.

Then every NFA for $L(A)$ has at least m states.

Proof.

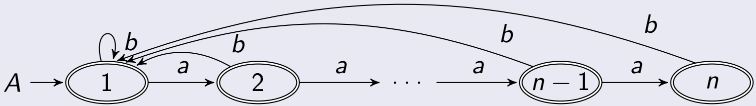
- X_i is reachable $\Rightarrow$ there is a string x_i such that $s \xrightarrow{x_i} X_i$
- Y_i is co-reachable $\Rightarrow$ there is a string y_i such that $Y_i \xrightarrow{y_i} acc$
- (2) and (3) $\Rightarrow X_i \cap Y_i = \{i\} \Rightarrow x_i y_i \in L(A)$
- $i > j$ and (3) $\Rightarrow X_i \cap Y_j = \emptyset \Rightarrow x_i y_j \notin L(A)$

$\Rightarrow$ the set $\{(x_i, y_i) \mid i = 1, 2, \ldots, m\}$ is a fooling set for $L(A)$, so every NFA for $L(A)$ has at least m states □

The Most Interesting Result of This Paper

Theorem 5 (4).

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.



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

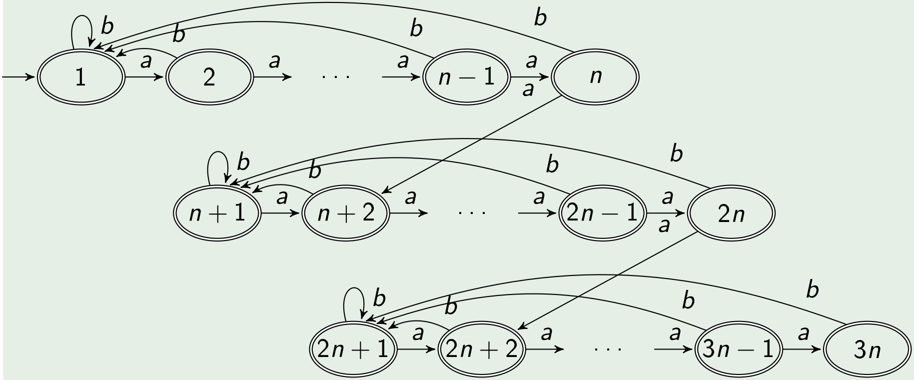
- the minimal partial DFA D for L^k has kn states and by Lemma 3, it is a minimal NFA for L^k
  - we show that in D, for every i with $1 \le i \le kn$, every set $\{i\}$ is reachable, and every set $\{1, 2, \ldots, i\}$ is co-reachable
  - using these pairs, we get a fooling set for L^k of size kn
- language L works also as a witness for concatenation

□

The Most Interesting Result of This Paper

Example ($k = 3$)

The minimal partial DFA D for L^3



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 – NSC on Subclasses of Convex Languages

	$K \cap L$	$K \cup L$	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	✓ ✓

Open Problems

From this paper

- Complexity of L^k on binary subword-closed and subword-convex languages

From our older papers

- Complexity of L^c: unknown on factor-convex and subword-convex
- Smaller alphabets ?
  - L^c all-sided ideal, subword-free, subword-closed, suffix-convex
  - KL
  - L^R subword-closed, subword-convex

ありがとう

Danke	Ďakujem	Kiitos
Paldies	Köszönöm	ขอบคุณ
Спаси́бо	Obrigado	Grazie

Summary and Open Problems

	L^k	$\lvert\Sigma\rvert$	L^+	$\lvert\Sigma\rvert$
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

Open problems

- Complexity of L^c: factor-convex and subword-convex
- Smaller alphabets:
  - L^c all-sided ideal, subword-free, subword-closed, suffix-convex
  - KL
  - L^k
  - L^R subword-closed, subword-convex