The Complexity of Operations on Union-Free and Star-Free Languages

Michal Hospodár

Mathematical Institute,
Slovak Academy of Sciences,
Košice, Slovakia

NCMA 2016, Debrecen, Hungary

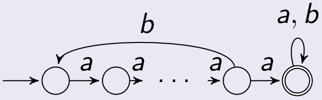
Subclasses of Regular Languages

Union-Free Languages

Regular expressions without $+$

Automata: one-cycle-free-path NFAs

(proven: Nagy, 2006)



Complexity of operations:

Jirásková, Masopust, DLT 2010

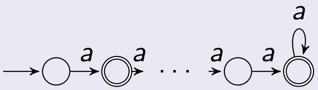
"...the case of unary union-free languages, is of interest too."

Star-Free Languages

Closed under Boolean operations and concatenation

Automata: aperiodic DFAs

(proven: Schützenberger, 1965)



Complexity of operations:

Holzer, Kutrib, Meckel, CIAA 2011: nondeterministic complexity (nsc)

Brzozowski, Liu, 2014: quotient complexity (equal to sc)

Known Complexities on General Alphabet

		Regular	$\lvert\Sigma\rvert$
$\mathrm{nsc}(K \cap L)$	Holzer	mn	2
$\mathrm{nsc}(K \cup L)$	Kutrib	$m+n+1$	2
$\mathrm{nsc}(KL)$	2003	$m+n$	2
$\mathrm{nsc}(L^2)$		$2n$	2
$\mathrm{nsc}(L^+)$		n	1
$\mathrm{nsc}(L^*)$	Rampersad	$n+1$	1
$\mathrm{nsc}(L^R)$	2006	$n+1$	2
$\mathrm{nsc}(L^c)$		2^n	2
$\mathrm{sc}(K \cap L)$		mn	2
$\mathrm{sc}(K \cup L)$	Yu	mn	2
$\mathrm{sc}(KL)$	Zhuang	$m2^n - 2^{n-1}$	2
$\mathrm{sc}(L^2)$	Salomaa	$n2^n - 2^{n-1}$	2
$\mathrm{sc}(L^*)$	1994	$2^{n-1} + 2^{n-2}$	2
$\mathrm{sc}(L^R)$		2^n	2

Known Complexities on General Alphabet

	Union-free	$\lvert\Sigma\rvert$	Regular	$\lvert\Sigma\rvert$
$\mathrm{nsc}(K \cap L)$	mn	2	mn	2
$\mathrm{nsc}(K \cup L)$	$m+n+1$	2	$m+n+1$	2
$\mathrm{nsc}(KL)$	$m+n$	2	$m+n$	2
$\mathrm{nsc}(L^2)$	$2n$	2	$2n$	2
$\mathrm{nsc}(L^+)$	n	1	n	1
$\mathrm{nsc}(L^*)$	$n+1$	1	$n+1$	1
$\mathrm{nsc}(L^R)$	n	1	$n+1$	2
$\mathrm{nsc}(L^c)$	2^n	3	2^n	2
$\mathrm{sc}(K \cap L)$	mn	2	mn	2
$\mathrm{sc}(K \cup L)$	mn	2	mn	2
$\mathrm{sc}(KL)$	$m2^n - 2^{n-1}$	2	$m2^n - 2^{n-1}$	2
$\mathrm{sc}(L^2)$	$n2^n - 2^{n-1}$	2	$n2^n - 2^{n-1}$	2
$\mathrm{sc}(L^*)$	$2^{n-1} + 2^{n-2}$	2	$2^{n-1} + 2^{n-2}$	2
$\mathrm{sc}(L^R)$	2^n	2	2^n	2

Jirásková
Masopust
2010

Jirásková
Šebej
DCFS 2011

Known Complexities on General Alphabet

	Star-free	$\|\Sigma\|$	Regular	$\|\Sigma\|$
$\mathrm{nsc}(K \cap L)$	mn	2	mn	2
$\mathrm{nsc}(K \cup L)$	$m+n+1$	2	$m+n+1$	2
$\mathrm{nsc}(KL)$	$m+n$	2	$m+n$	2
$\mathrm{nsc}(L^2)$	$2n$	2	$2n$	2
$\mathrm{nsc}(L^+)$	n	1	n	1
$\mathrm{nsc}(L^*)$	$n+1$	2	$n+1$	1
$\mathrm{nsc}(L^R)$	$n+1$	2	$n+1$	2
$\mathrm{nsc}(L^c)$	2^n	2	2^n	2
$\mathrm{sc}(K \cap L)$	mn	2	mn	2
$\mathrm{sc}(K \cup L)$	mn	2	mn	2
$\mathrm{sc}(KL)$	$m2^n - 2^{n-1}$	4	$m2^n - 2^{n-1}$	2
$\mathrm{sc}(L^2)$			$n2^n - 2^{n-1}$	2
$\mathrm{sc}(L^*)$	$2^{n-1} + 2^{n-2}$	4	$2^{n-1} + 2^{n-2}$	2
$\mathrm{sc}(L^R)$	$2^n - 1$	$n-1$	2^n	2

Holzer
Kutrib
Meckel
2011
Domaratzki
Okhotin
2009

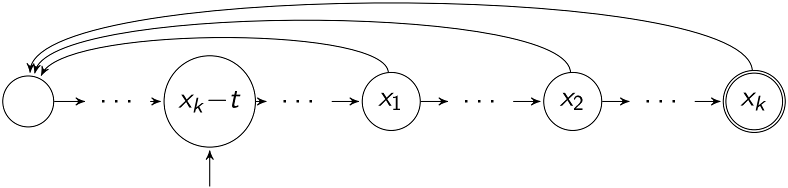
Brzozowski
Liu
2014

Known Complexities on Unary Alphabet

	Unary regular	Source
$\mathrm{nsc}(K \cap L)$	mn ; $\gcd(m,n) = 1$	Holzer, Kutrib 2003, Thm. 4
$\mathrm{nsc}(K \cup L)$	$m + n + 1$; $m \neq kn$	Holzer, Kutrib 2003, Thm. 2
$\mathrm{nsc}(KL)$	$\geq m + n - 1$, $\leq m + n$	Holzer, Kutrib 2003, Thm. 8
$\mathrm{nsc}(L^2)$	$\geq 2n - 1$, $\leq 2n$	implies from HK 2003, Thm. 8
$\mathrm{nsc}(L^+)$	n	Holzer, Kutrib 2003, Thm. 9
$\mathrm{nsc}(L^*)$	$n + 1$	Holzer, Kutrib 2003, Thm. 9
$\mathrm{nsc}(L^c)$	$2^{\Theta(\sqrt{n \log n})}$	Chrobak 1986, Thm. 4.4, 4.5
$\mathrm{sc}(K \cap L)$	mn ; $\gcd(m,n) = 1$	Pighizzini, Shallit 2002
$\mathrm{sc}(K \cup L)$	mn ; $\gcd(m,n) = 1$	Pighizzini, Shallit 2002
$\mathrm{sc}(KL)$	mn ; $\gcd(m,n) = 1$	YZS 1994, Thm. 5.4, 5.5
$\mathrm{sc}(L^2)$	$2n - 1$	Rampersad 2006, Thm. 2??
$\mathrm{sc}(L^*)$	$(n-1)^2 + 1$	YZS 1994, Thm. 5.3

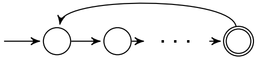
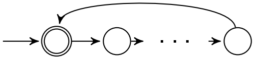
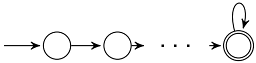
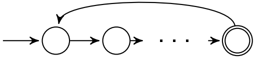
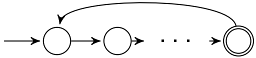
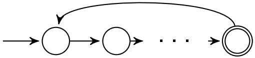
One cycle-free path $\Rightarrow$ single final state, no concurrent paths
Minimal NFA – congruent states differ in finality
Corollary: cycles are overlapping as much as possible

A minimal NFA for language $L = a^t(a^{x_1})^*(a^{x_2})^* \dots (a^{x_k})^*$



$k = 0 \Rightarrow$ finite language
$k \geq 2$, x_1 and x_2 are co-prime $\Rightarrow$ co-finite language

Worst-Case Examples: Unary Union-Free Languages

Operation	NSC	SC
		
Intersection	HK 2003, Thm. 4	Pighizzini, Shallit 2002
Union	HK 2003, Thm. 2	
		
Concatenation	HK 2003, Thm. 8	YZS 1994, Thm. 5.5
Square	H. 2016 – upper bound	Domaratzki, Okhotin 2009
		
Positive closure	HK 2003, Thm. 9	
Kleene star	HK 2003, Thm. 9	YZS 1994, Thm. 5.3

Complementation of Unary Union-Free Languages

Upper Bound

Minimal NFA for Unary Union-Free Language

If x_1, x_2 are co-prime, then L is co-finite.

We look for the longest string in L^c

The biggest L^c is obtained if $x_1 = n-1$ and $x_2 = n$

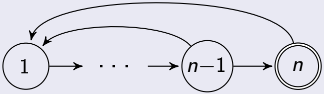
String in L ... number $i(n-1) + jn$ for $i \geq 1$, $j \geq 0$

Longest string in L^c ... $(n-1)^2 - 1$

Upper bound of nsc ... $(n-1)^2$

Lower Bound

$L = a^{n-1}(a^{n-1})^*(a^n)^*$



$\mathrm{nsc}(L^c) = (n-1)^2$

Unary Star-Free Languages

General alphabet: finite $\cup$ co-finite $\subset$ star-free

Unary alphabet: finite $\cup$ co-finite $=$ star-free

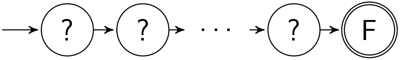
Unary convex $\subset$ unary star-free

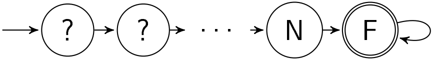
Finite $L \subseteq \{a^i \mid 0 \le i \le n-1\}$

longest accepted string is a^{n-1}

Co-finite $L \supseteq a^{n-1}a^*$

longest rejected string is a^{n-2}





Worst-Case Examples: Unary Star-Free Languages

Operation	NSC witness	SC witness
Intersection	co-finite	co-finite
Union	finite	finite
	H. 2016	BL 2014, Thm. 6
	co-finite	co-finite
Concatenation	H. 2016	BL 2014, Thm. 6
Square	H. 2016	Rampersad 2006, Thm. 2??
	co-finite	Finite languages:
Positive closure	H. 2016	Câmpeanu, Culik,
Kleene star	H. 2016	Salomaa, Yu, WIA 2001

Complementation of Unary Star-Free Languages

NFA-to-Chrobak transformation

Input: n-state NFA
Output: Chrobak normal form
tail: $\leq n^2 - 2$ states
cycles: $\leq n - 1$ states
(Geffert 2007, Thm. 3.5).

Corollary

Since unary aperiodic NFAs have no cycles but loops, such NFA in Chrobak normal form has at most $n^2 - 1$ states.

This NFA is also a complete DFA: upper bound $O(n^2)$ on complementation of co-finite languages.

Lower Bound

Jirásková, Mlynárčik, DCFS 2014:
finite language $L = \{a^{n-1}\}$ with $\mathrm{nsc}(L^c) = \Theta(\sqrt{n})$
$\Rightarrow$ complement of co-finite languages $= \Theta(n^2)$

Obrazok Chrobakovho tvaru asi dat?

Summary - State Complexity of Operations on Union-Free Languages and Star-Free Languages

	Unary union-free	Unary star-free	Unary regular
$\mathrm{nsc}(K \cap L)$	mn ; $\gcd(m,n) = 1$	$\max\{m,n\}$	mn ; $\gcd(m,n) = 1$
$\mathrm{nsc}(K \cup L)$	$m + n + 1$; $m \neq kn$	$\max\{m,n\}$	$m + n + 1$; $m \neq kn$
$\mathrm{nsc}(KL)$	$m + n - 1$	$m + n - 1$	$\geq m + n - 1$
$\mathrm{nsc}(L^2)$	$2n - 1$	$2n - 1$	$\geq 2n - 1$
$\mathrm{nsc}(L^+)$	n	n	n
$\mathrm{nsc}(L^*)$	$n + 1$	n	$n + 1$
$\mathrm{nsc}(L^c)$	$(n-1)^2$	$\Theta(n^2)$	$2^{\Theta(\sqrt{n \log n})}$
$\mathrm{sc}(K \cap L)$	mn ; $\gcd(m,n) = 1$	$\max\{m,n\}$	mn ; $\gcd(m,n) = 1$
$\mathrm{sc}(K \cup L)$	mn ; $\gcd(m,n) = 1$	$\max\{m,n\}$	mn ; $\gcd(m,n) = 1$
$\mathrm{sc}(KL)$	mn ; $\gcd(m,n) = 1$	$m + n - 1$	mn ; $\gcd(m,n) = 1$
$\mathrm{sc}(L^2)$	$2n - 1$	$2n - 1$	$2n - 1$
$\mathrm{sc}(L^*)$	$(n-1)^2 + 1$	$n^2 - 7n + 13$	$(n-1)^2 + 1$

Summary - State Complexity of Operations on Union-Free Languages and Star-Free Languages

	Union-free	$\|\Sigma\|$	Star-free	$\|\Sigma\|$	Regular	$\|\Sigma\|$
$\mathrm{nsc}(K \cap L)$	mn	2	mn	2	mn	2
$\mathrm{nsc}(K \cup L)$	$m+n+1$	2	$m+n+1$	2	$m+n+1$	2
$\mathrm{nsc}(KL)$	$m+n$	2	$m+n$	2	$m+n$	2
$\mathrm{nsc}(L^2)$	$2n$	2	$2n$	2	$2n$	2
$\mathrm{nsc}(L^+)$	n	1	n	1	n	1
$\mathrm{nsc}(L^*)$	$n+1$	1	$n+1$	2	$n+1$	1
$\mathrm{nsc}(L^R)$	n	1	$n+1$	2	$n+1$	2
$\mathrm{nsc}(L^c)$	2^n	3	2^n	2	2^n	2
$\mathrm{sc}(K \cap L)$	mn	2	mn	2	mn	2
$\mathrm{sc}(K \cup L)$	mn	2	mn	2	mn	2
$\mathrm{sc}(KL)$	$m2^n - 2^{n-1}$	2	$m2^n - 2^{n-1}$	4	$m2^n - 2^{n-1}$	2
$\mathrm{sc}(L^2)$	$n2^n - 2^{n-1}$	2	?	?	$n2^n - 2^{n-1}$	2
$\mathrm{sc}(L^*)$	$2^{n-1} + 2^{n-2}$	2	$2^{n-1} + 2^{n-2}$	4	$2^{n-1} + 2^{n-2}$	2
$\mathrm{sc}(L^R)$	2^n	2	$2^n - 1$	$n-1$	2^n	2

Thank You For Attention

Köszonöm a figyelemért