The State Complexity of Permutations on Finite Languages Over Binary Alphabets

Alexandros Palioudakis[1] Da-Jung Cho[1] Daniel Goč[2]
Yo-Sub Han[1] Sang-Ki Ko[1] Kai Salomaa[2]

[1]Department of Computer Science
Yonsei University

[2]School of Computing
Queen's University

Descriptional Complexity of Formal Systems, 2015

Outline

DCFS 2014, Turku, Finland

- Cho et al., *Pseudo-Inversion on Formal Languages*, UCNC 2014
- Cho et al., *State Complexity of Inversion Operations*, DCFS 2014

DCFS 2014, Turku, Finland

We proved the following theorem:

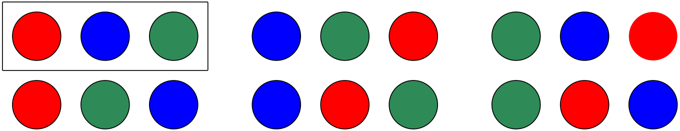
Theorem

Given a string w over Σ, $\mathbb{PI}^*(w) = \pi(w)$, where $\pi(w)$ is the set of all permutations of w.

What is the property of permutation operation?

Permutation

- Permutation is well-known in combinatorics.



An example of different permutations of three distinct balls

- In formal language theory, permutation per on a string w is

$$\mathrm{per}(w) = \{u \in \Sigma^* \mid |w|_a = |u|_a, \text{ for } a \in \Sigma\}.$$

Permutation

- Regular languages are not closed under permutation.

Example

Let a language $L = L(a \cdot b)^*$. Then,

$$\mathrm{per}(L) = \{w \mid |w|_a = |w|_b\}.$$

State Complexity

- Deterministic State Complexity (sc): the number of states in a minimal DFA.
- Nondeterministic State Complexity (nsc): the number of states in an NFA.

- State complexity of some operations:
  - $\mathrm{sc}(L_1) = m$ and $\mathrm{sc}(L_2) = n$
  - $\mathrm{sc}(L_1 \cap L_2) = mn$
  - $\mathrm{sc}(L_1 \cup L_2) = mn$
  - $\mathrm{sc}(L_1)^R = 2^m$

State Complexity of Permutations on Finite Languages

- Upper bound construction for per of a binary finite language with restriction.
  - restriction: length of strings s.t. $L \subseteq \{a, b\}^{n-1}$
- State complexity of per on a binary finite language.

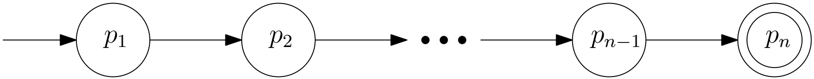
Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(L))$

Lemma

Let n be a positive integer and $L \subseteq \{a, b\}^{n-1}$ be a finite language such that $\mathrm{sc}(L) = n$. Then, we have the following inequality for the state complexity of the permutation of L:

$$\mathrm{sc}(\mathrm{per}(L)) \leq \frac{n^2+n+1}{3}$$

Let A be the minimal DFA for $L \subseteq \{a, b\}^{n-1}$.



A DFA A forms a chain.

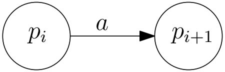
Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(L))$

Lemma

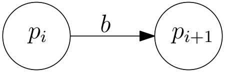
Let n be a positive integer and $L \subseteq \{a, b\}^{n-1}$ be a finite language such that $\mathrm{sc}(L) = n$. Then, we have the following inequality for the state complexity of the permutation of L:

$$\mathrm{sc}(\mathrm{per}(L)) \leq \frac{n^2+n+1}{3}$$

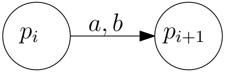
Let A be the minimal DFA for $L \subseteq \{a, b\}^{n-1}$.



(1) a-transition



(2) b-transition



(3) $a\&b$-transition

Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(L))$

Lemma

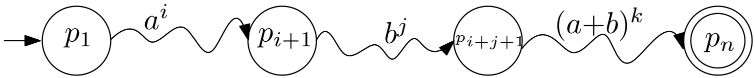
Let n be a positive integer and $L \subseteq \{a, b\}^{n-1}$ be a finite language such that $\mathrm{sc}(L) = n$. Then, we have the following inequality for the state complexity of the permutation of L:

$$\mathrm{sc}(\mathrm{per}(L)) \leq \frac{n^2+n+1}{3}$$

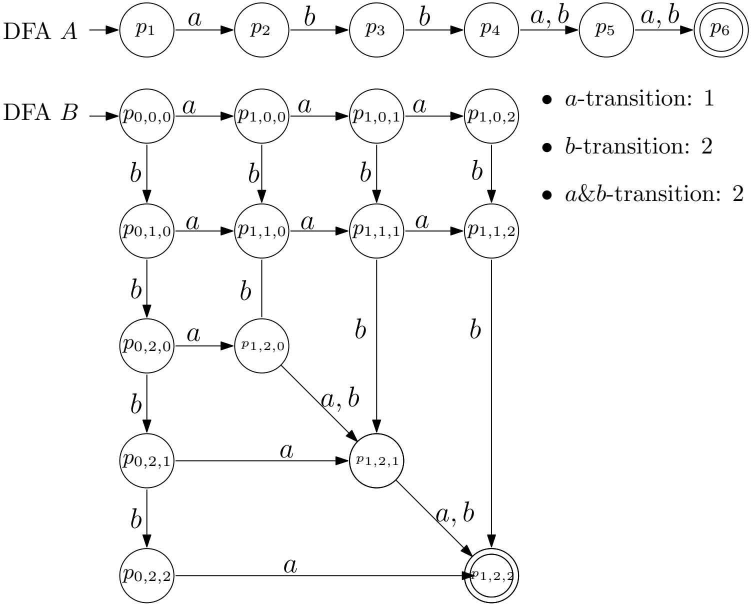
The order of a, b and $a\&b$-transitions does not affect $\mathrm{per}(L)$.

- Assume that for $i, j, k \geq 1$ and $i+j+k = n-1$, L is of the form

$$a^i b^j (a+b)^k.$$

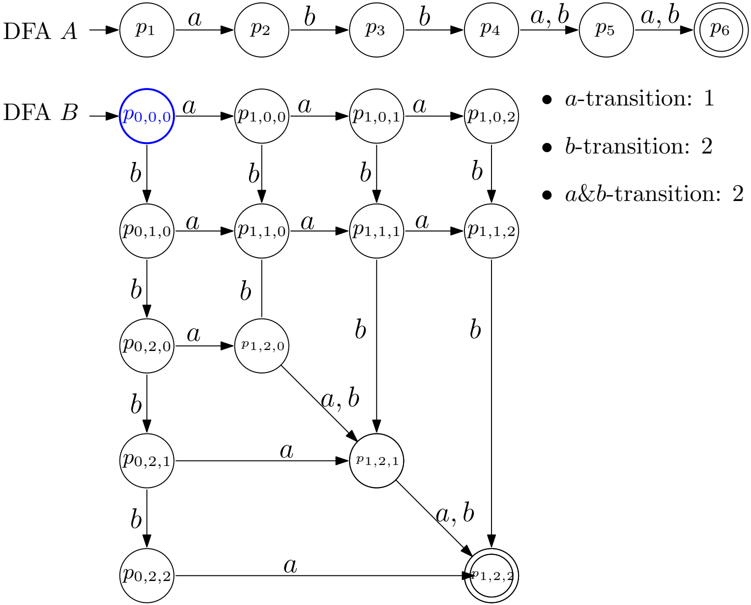


Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(ab^2(a+b)^2))$



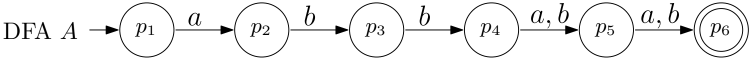
- a-transition: 1
- b-transition: 2
- $a\&b$-transition: 2

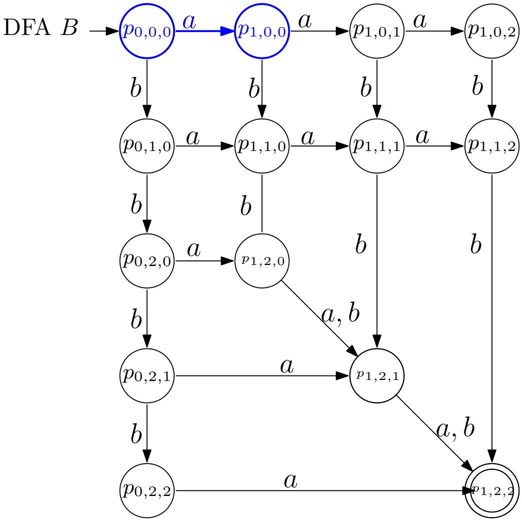
Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(ab^2(a+b)^2))$



- a-transition: 1
- b-transition: 2
- $a\&b$-transition: 2

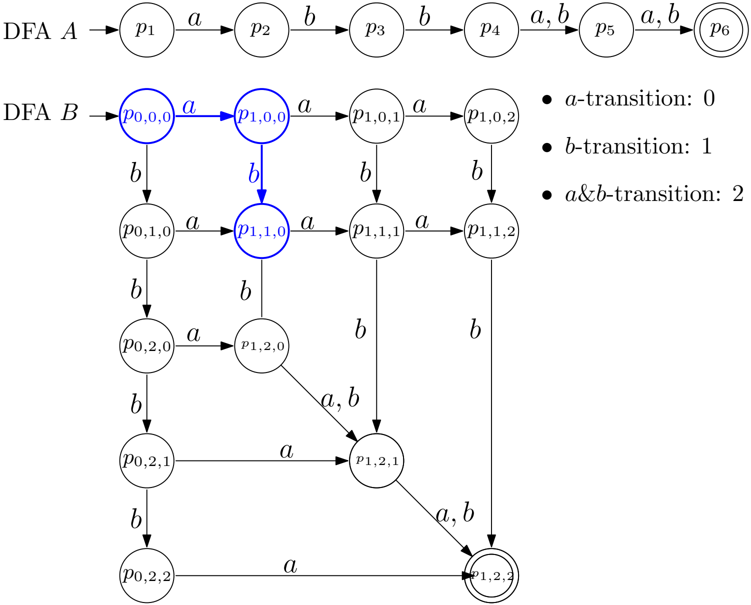
Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(ab^2(a+b)^2))$





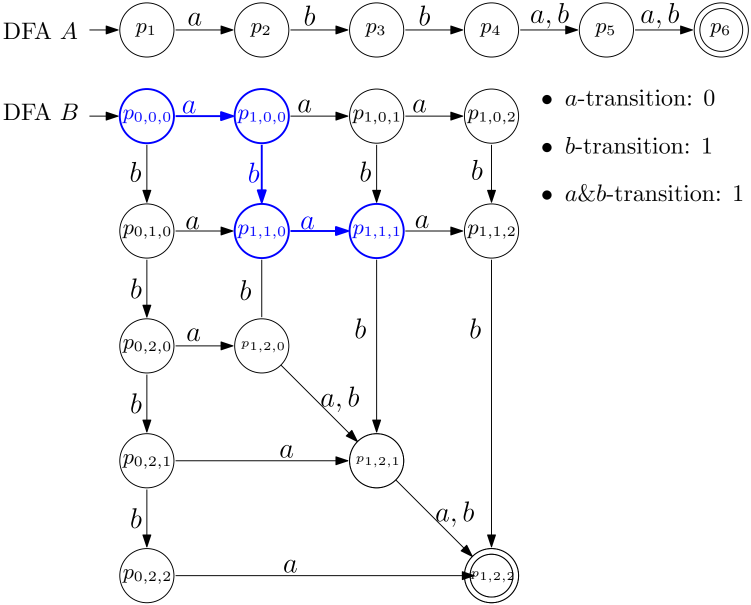
- a-transition: 0
- b-transition: 2
- $a\&b$-transition: 2

Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(ab^2(a+b)^2))$



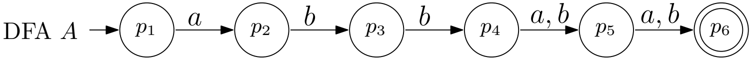
- a-transition: 0
- b-transition: 1
- $a\&b$-transition: 2

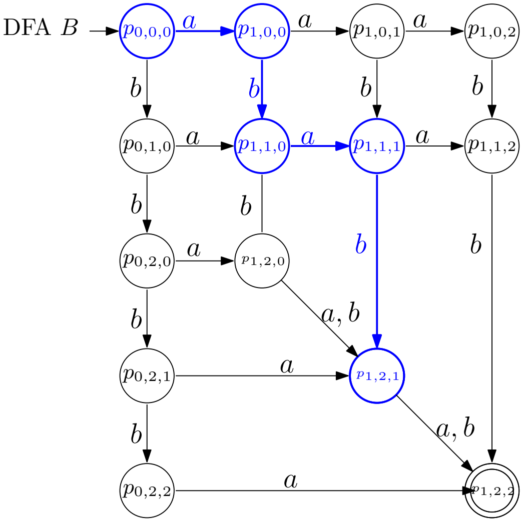
Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(ab^2(a+b)^2))$



- a-transition: 0
- b-transition: 1
- $a\&b$-transition: 1

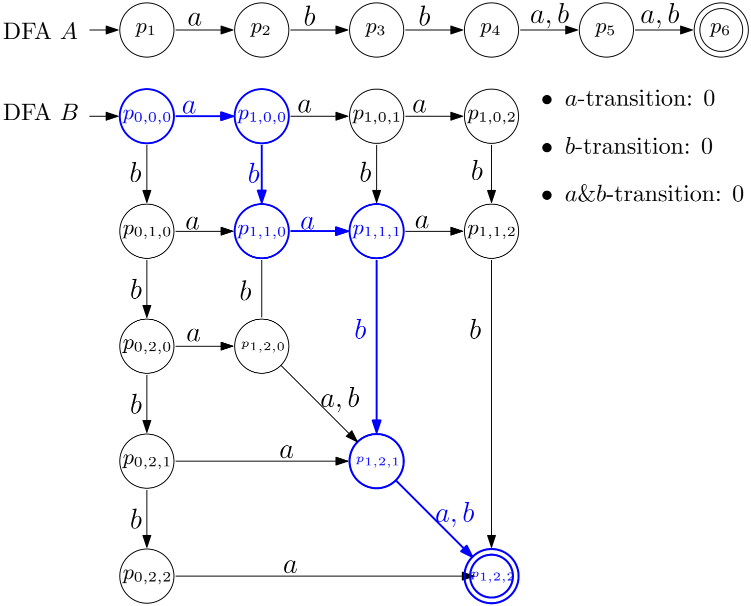
Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(ab^2(a+b)^2))$





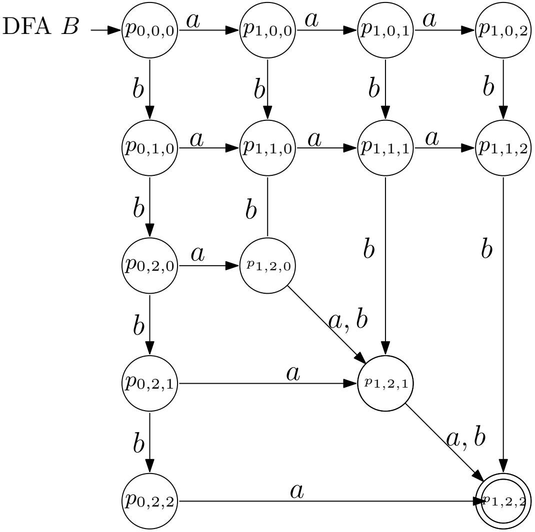
- a-transition: 0
- b-transition: 0
- $a\&b$-transition: 1

Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(ab^2(a+b)^2))$



- a-transition: 0
- b-transition: 0
- $a\&b$-transition: 0

Upper Bound Construction for $\text{sc}(\text{per}(ab^2(a+b)^2))$



$(i+1) \cdot (j+1) + (k \cdot i) + (k \cdot j) + k$ states

Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(L))$

Lemma

Let n be a positive integer and $L \subseteq \{a, b\}^{n-1}$ be a finite language such that $\mathrm{sc}(L) = n$. Then, we have the following inequality for the state complexity of the permutation of L:

$$\mathrm{sc}(\mathrm{per}(L)) \leq \frac{n^2+n+1}{3}$$

We can construct a DFA B for $\mathrm{per}(L(A))$ with

$$(i+1) \cdot (j+1) + (k \cdot i) + (k \cdot j) + k \text{ states} .$$

Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(L))$

Lemma

Let n be a positive integer and $L \subseteq \{a, b\}^{n-1}$ be a finite language such that $\mathrm{sc}(L) = n$. Then, we have the following inequality for the state complexity of the permutation of L:

$$\mathrm{sc}(\mathrm{per}(L)) \leq \frac{n^2+n+1}{3}$$

- Let function $f(i, j, k) = (i+1) \cdot (j+1) + (k \cdot i) + (k \cdot j) + k$.
- Then, f is maximized when $i = j = k = \frac{n-1}{3}$.

$$\max f(i, j, k) = \begin{cases} \frac{n^2+n+1}{3} & \text{if } n \equiv 1 \pmod 3 \\ \frac{n^2+n}{3} & \text{otherwise .} \end{cases}$$

Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(L))$

Lemma

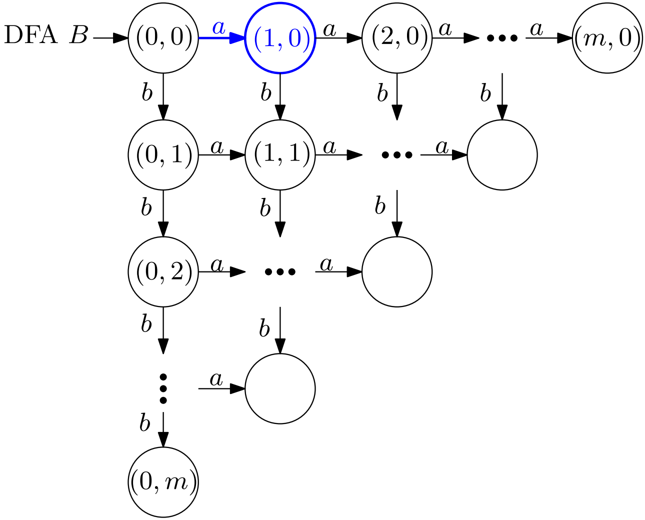
Let L be a binary finite language and $m = \max\{|w| \mid w \in L\}$ *for some positive integer m. Then, we have* $\mathrm{sc}(\mathrm{per}(L)) \leq \frac{m^2+m+2}{2}$.

Given a DFA A for L over the binary alphabet,

- We construct a DFA B for $\mathrm{per}(L(A))$.
- B has states of the form (i, j), for $w \in L(A)$:
  - i tracks $|w|_a$,
  - j tracks $|w|_b$.
- For all states (i, j), $i + j \leq m$.

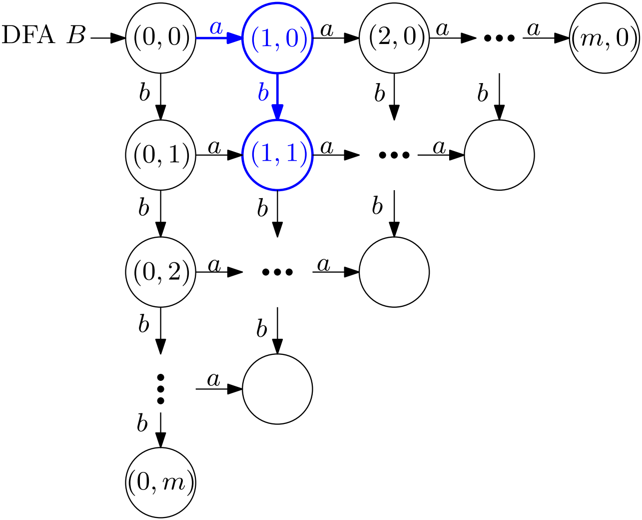
Upper Bound Construction for sc(per(L))

- For (i, j), i tracks $|w|_a$ and j tracks $|w|_b$.
- For (i, j), $i + j \leq m$.



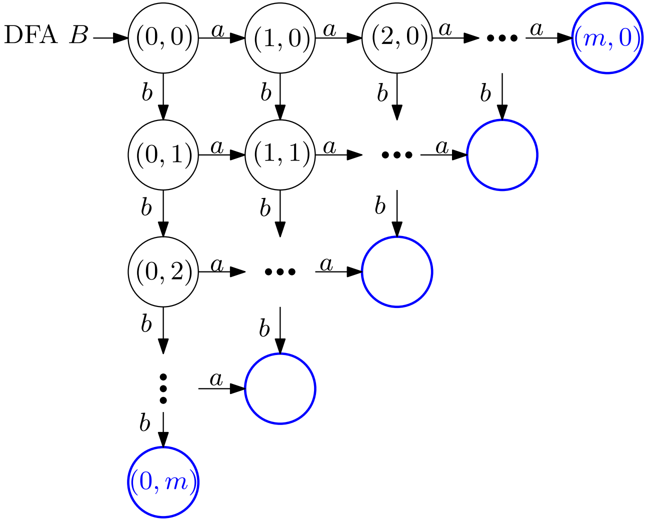
Upper Bound Construction for sc(per(L))

- For (i, j), i tracks $|w|_a$ and j tracks $|w|_b$.
- For (i, j), $i + j \leq m$.



Upper Bound Construction for $\text{sc}(\text{per}(L))$

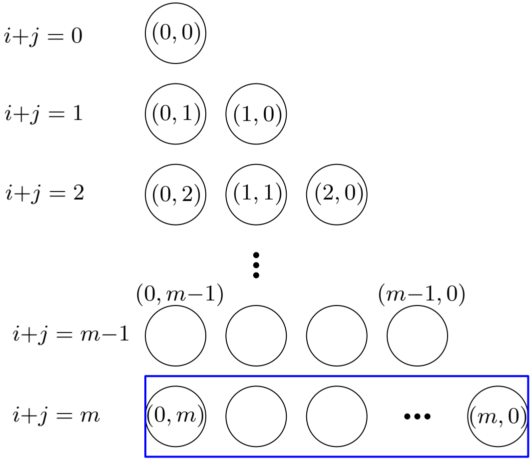
- For (i,j), i tracks $|w|_a$ and j tracks $|w|_b$.
- For (i,j), $i+j \leq m$.



Upper Bound Construction for sc(per(L))

The total number of states of B for per($L(A)$):

$$1 + 2 + \cdots + m + 1 = \frac{m \cdot (m+1)}{2} + 1$$



Upper Bound Construction for $\mathrm{sc}(\mathrm{per}(L))$

Corollary

Let L be a binary finite languages and $\mathrm{sc}(L) = n$ for some positive integer n. Then, we have

$$\mathrm{sc}(\mathrm{per}(L)) \leq \frac{n^2 - n + 2}{2}.$$

Since $1 + \max\{|w| \mid w \in L\} \leq \mathrm{sc}(L)$,

$$m \leq \mathrm{sc}(L) - 1 = n - 1.$$

Then,

$$\mathrm{sc}(\mathrm{per}(L)) \leq \frac{m^2 + m + 2}{2} = \frac{(n-1)^2 + (n-1) + 2}{2}.$$

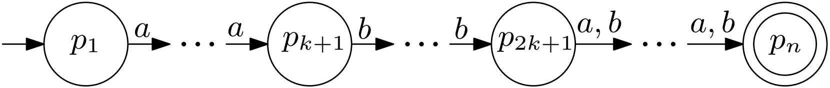
Lower Bound for $\mathrm{sc}(\mathrm{per}(L))$

Theorem

For any $n_0 \in \mathbb{N}$, there exists a regular language L with $\mathrm{sc}(L) = n$, for $n \geq n_0$, such that

$$\mathrm{sc}(\mathrm{per}(L)) \geq \frac{n^2 + n + 1}{3}.$$

Let $n = 3k + 1 \geq n_0, k \in \mathbb{N}$ and $L_n = L(a^k b^k (a + b)^k)$.



DFA A for L_n

Lower Bound for $\mathrm{sc}(\mathrm{per}(L))$

Let $n = 3k + 1 \geq n_0, k \in \mathbb{N}$ and $L_n = L(a^k b^k (a + b)^k)$.
Then,

$$\mathrm{per}(L_n) = \{w \in \Sigma^{3 \cdot k} \mid |w|_a, |w|_b \geq k \text{ and } n = 3 \cdot k{+}1\}.$$

Let X and Y be the sets of strings:

- $X = \{a^i b^j \mid 0 \leq i \leq 2k \text{ and } 1 \leq j \leq k\}$,
- $Y = \{a^i b^j \mid 0 \leq i < k \text{ and } k < j \leq 2k\}$.

Inequivalent check

All strings of $X \cup Y$ are pairwise inequivalent with respect to the Myhill-Nerode congruence of $\mathrm{per}(L_n)$.

Lower Bound for $\mathrm{sc}(\mathrm{per}(L))$

Let $u = a^i b^j$ and $u' = a^{i'} b^{j'}$ be two arbitrary distinct strings $\in X \cup Y$.

(i) $u, u' \in X$

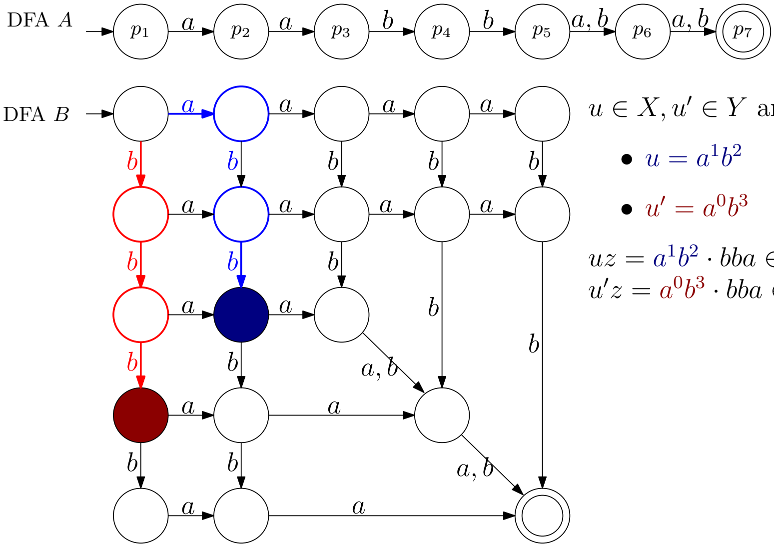
(ii) $u, u' \in Y$

(iii) $u \in X$ and $u' \in Y$

Case (iii)

- Let $u \in X, u' \in Y$ and $|u| = |u'|$.
- Because of following condition:
  - $X = \{a^i b^j \mid 0 \leq i \leq 2k \text{ and } 1 \leq j \leq k\}$
  - $Y = \{a^i b^j \mid 0 \leq i < k \text{ and } k < j \leq 2k\}$
- $|u|_b \leq k$ and $|u'|_b > k$.
- Then, $|u|_a > |u'|_a$.
- For $z = a^{k-i} b^{2k-j}$, $uz = a^i b^j \cdot a^{k-i} b^{2k-j} \in \mathrm{per}(L_n)$.
- But, $u'z = a^{i'} b^{j'} \cdot a^{k-i} b^{2k-j} \notin \mathrm{per}(L_n)$.

Lower Bound for $\mathrm{sc}(\mathrm{per}(a^2b^2(a+b)^2))$

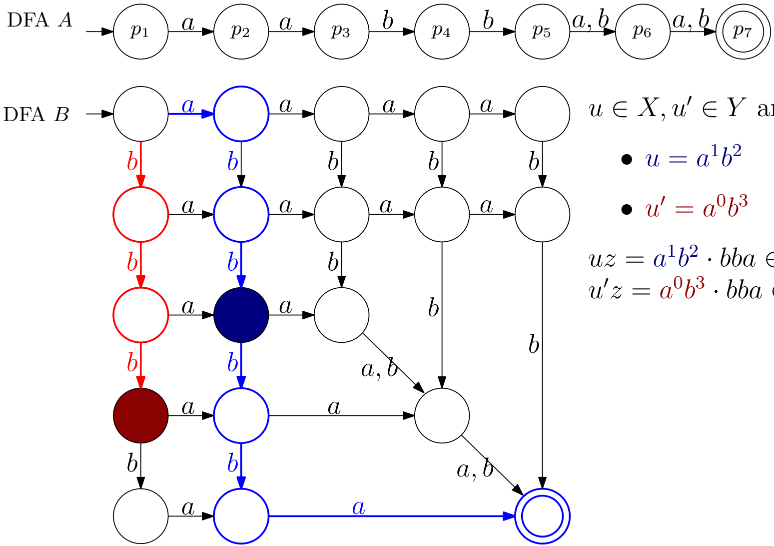


$u \in X, u' \in Y$ and $|u| = |u'|$

- $u = a^1b^2$
- $u' = a^0b^3$

$uz = a^1b^2 \cdot bba \in \mathrm{per}(L(A))$
$u'z = a^0b^3 \cdot bba \notin \mathrm{per}(L(A))$

Lower Bound for $\mathrm{sc}(\mathrm{per}(a^2b^2(a+b)^2))$



$u \in X, u' \in Y$ and $|u| = |u'|$

- $u = a^1b^2$
- $u' = a^0b^3$

$uz = a^1b^2 \cdot bba \in \mathrm{per}(L(A))$
$u'z = a^0b^3 \cdot bba \notin \mathrm{per}(L(A))$

Lower Bound for $\mathrm{sc}(\mathrm{per}(L))$

Theorem

For any $n_0 \in \mathbb{N}$, there exists a regular language L with $\mathrm{sc}(L) = n$, for $n \geq n_0$, such that

$$\mathrm{sc}(\mathrm{per}(L)) \geq \frac{n^2 + n + 1}{3}.$$

- $X = \{a^i b^j \mid 0 \leq i \leq 2k \text{ and } 1 \leq j \leq k\}$,
- $Y = \{a^i b^j \mid 0 \leq i < k \text{ and } k < j \leq 2k\}$.

Thus, the $\#$ of states of the minimal DFA has at least

$$(2 \cdot k+1) \cdot (k+1) + k^2 = 3 \cdot k^2 + 3 \cdot k + 1 \text{ states .}$$

Since $n = 3 \cdot k + 1$, $k = \frac{n-1}{3}$,

$$3 \cdot (\frac{n-1}{3})^2 + 3 \cdot (\frac{n-1}{3}) + 1 = \frac{n^2 + n + 1}{3}$$

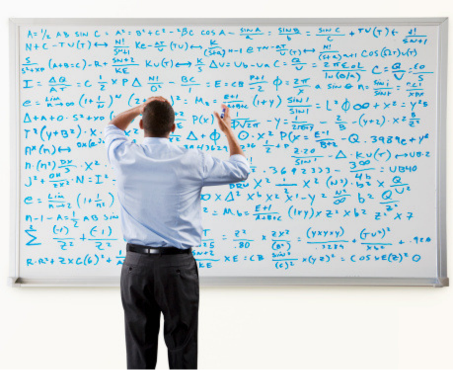
Summary

Deterministic state complexity

- Upper bound for $\mathrm{per}(L)$, where $L \subseteq \{a, b\}^{n-1}$: $\frac{n^2+n+1}{3}$
- Upper bound for $\mathrm{per}(L)$: $\frac{n^2-n+2}{2}$
- Lower bound for $\mathrm{per}(L)$: $\frac{n^2+n+1}{3}$

Open Problems

- State complexity of permutation over non-binary languages
- State complexity of permutation of set of equal length strings
- Nondeterministic state complexity of permutation on finite languages



Thank you!

WELCOME TO CIAA 2016

SEOUL, SOUTH KOREA

JULY 19 -- 22, 2016