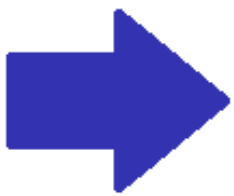
Operations on Permutation Automata

Michal Hospodár

Mathematical Institute, Slovak Academy of Sciences
Košice, Slovakia

Joint work with Peter Mlynárčik
Submitted to DLT 2020
Presented online on August 16, 2021



Finite Automata

A nondeterministic finite automaton with multiple initial states (MNFA) is a quintuple $N = (Q, \Sigma, \cdot, I, F)$ where

- Q is the set of states,
- Σ is the input alphabet,
- I is the set of initial states,
- F is the set of final states,
- $\cdot$ is the transition function from $Q \times \Sigma$ to 2^Q which can be extended to the domain $2^Q \times \Sigma^*$.

A complete DFA is a quintuple $A = (Q, \Sigma, \cdot, s, F)$ where

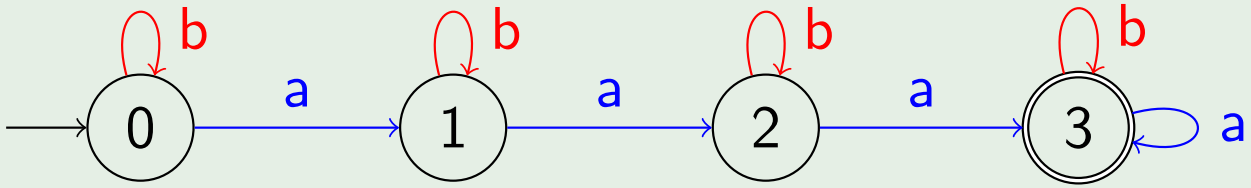
- s is the single initial state, and
- $\cdot$ is a function from $Q \times \Sigma$ to Q (extended to domain $Q \times \Sigma^*$).

The language accepted by DFA A: $L(A) = \{w \in \Sigma^* \mid s \cdot w \in F\}$

Transformations in DFAs

In DFAs, each symbol σ in Σ induces a transformation on Q given by $q \mapsto q \cdot \sigma$.

Example: DFA $(\{0,1,2,3\},\{a,b\},\cdot,0,\{3\})$ with $(0,1,2,3)\cdot a = (1,2,3,3)$ and $(0,1,2,3)\cdot b = (0,1,2,3)$



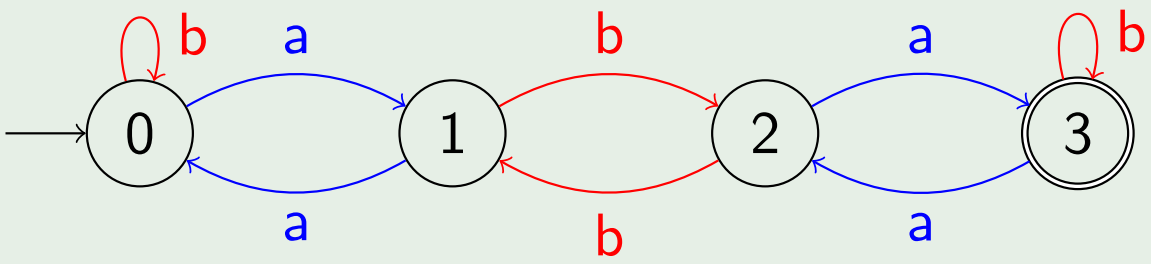
This DFA accepts the language $\{w \in \{a,b\}^* \mid |w|_a \geq 3\}$

Permutation DFAs

- If transformation induced by each symbol in Σ is a permutation, then we speak of a permutation automaton
- Languages accepted by permutation automata are called permutation languages or group languages
- Each permutation is a composition of several cycles
- An expression $a\colon (p_1, p_2, \ldots, p_k)\cdots(q_1, q_2, \ldots, q_\ell)$ denotes the permutation induced by symbol a:
  - all states in the cycles are permuted cyclically,
  - all other states are sent to themselves

Example

$a\colon (0,1)(2,3)$
$b\colon (1,2)$



Motivation and History

G. Thierrin: Permutation automata. Math. Syst. Theory 2(1), 83-90 (1968)

The class of group languages

- is closed under Boolean operations, reversal, and quotients
- is not closed under concatenation and Kleene star

Proofs are based on group properties of transition monoids of languages and cancellative congruences

This paper:

Operations on permutation automata. In: Proc. DLT 2020, pp. 122-136

- simpler proofs for closure and non-closure properties, based on DFAs
- more non-closure properties: L^k, L^+, $K \,!\, L$, $K \sqcup\!\sqcup L$, shift(L), per(L)
- state complexity of operations except for shuffle, shift, permutation

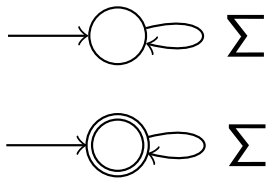
$K \,!\, L$ = cut operation, defined by Berglund et al. (2013)

Methods

Lemma

If L is a group language, then the minimal DFA for L is a permutation DFA.

Sink state:



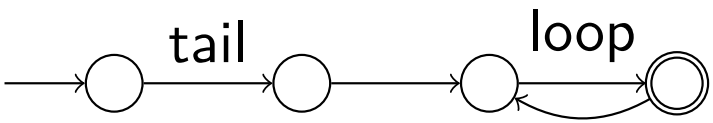
- Group language = its minimal DFA is a permutation DFA
- DFA with a reachable sink state is not a permutation DFA

Minimality of unary DFAs:

Lemma (Nicaud, 1999)

A unary DFA is minimal if and only if

- *its loop is minimal,*
- *the last state of tail and the last state of loop do not have the same finality.*



Minimality of binary DFAs:

Let $N = (Q, \Sigma, \cdot, I, F)$ be an MNFA

- $N^R = (Q, \Sigma, \cdot^R, F, I)$,
- $\mathcal{D}(N) = (2^Q, \Sigma, \cdot, I, F')$ where $F' = \{S \in 2^Q \mid S \cap F \neq \emptyset\}$

Lemma (Jirásková, Krajňáková, 2019)

If for an NFA N, each singleton set $\{q\}$ is reachable in N^R, then all states of $\mathcal{D}(N)$ are pairwise distinguishable.

Closure Properties

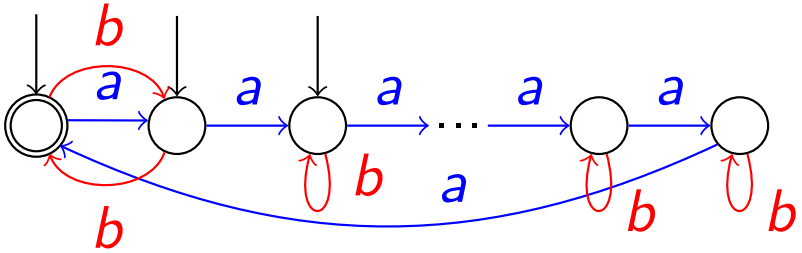
Definition

MDFA := MNFA with $|q \cdot \sigma| = 1$ for each $q \in Q$ and $\sigma \in \Sigma$

Let N be a permutation MDFA

Then N^R is a permutation MDFA and $\mathcal{D}(N)$ is a permutation DFA with $\binom{|Q|}{|I|}$ reachable states where I is the set of initial states of N

The hardest example: $\mathrm{sc} = \binom{n}{\lceil n/2 \rceil}$



(first $\lceil n/2 \rceil$ states are initial)

Let A and B be permutation DFAs with $L(A) = K$ and $L(B) = L$. Then

- DFAs for K^c and KL^{-1} have the same transitions as A;
- the product automaton $A \times B$ is a permutation DFA;
- K^R and $L^{-1}K$ are accepted by permutation MDFAs.

Theorem (cf. Thierrin, Section 4)

The class of group languages is closed under Boolean operations, reversal, and left and right quotient.

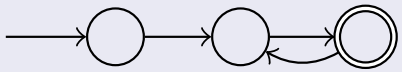
Non-closure Properties

Find group languages s.t. the result of an operation is not a group language:

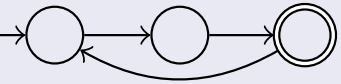
Square

A 

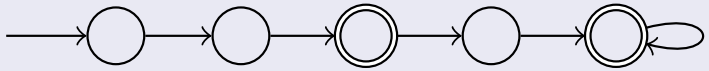
Minimal DFA for $L(A)^2$:



Star and Positive Closure

C 

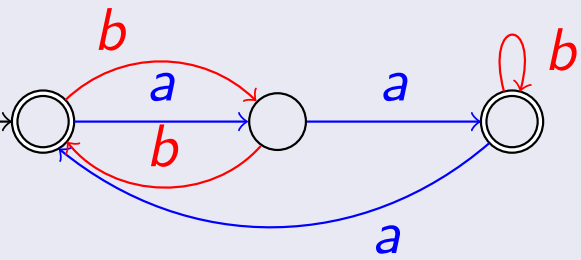
DFA for $L(C)^+$ has a reachable sink state:

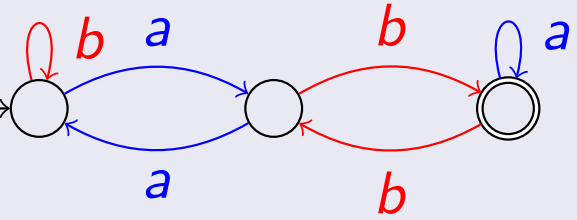


Concatenation, Power, Cut, Shuffle

Let $A = B$. Then
$L(A)L(B) =$
$L(A) \,!\, L(B) =$
$L(A) \sqcup\!\sqcup L(B) = L(A)^2$.

Cyclic Shift and Permutation

Cyclic Shift: D 

Permutation: E 

Upper Bounds for Concatenation and Square

Construct an NFA for concatenation or square in the usual way

Concatenation: $L(A)L(B)$
Regular upper bound: $m2^n - 2^{n-1}$

- If B is a permutation DFA, then from Q_B only Q_B is reachable

$\Rightarrow$ So for each p, q in Q_A, the sets $\{p\} \cup Q_B$ and $\{q\} \cup Q_B$ are equivalent

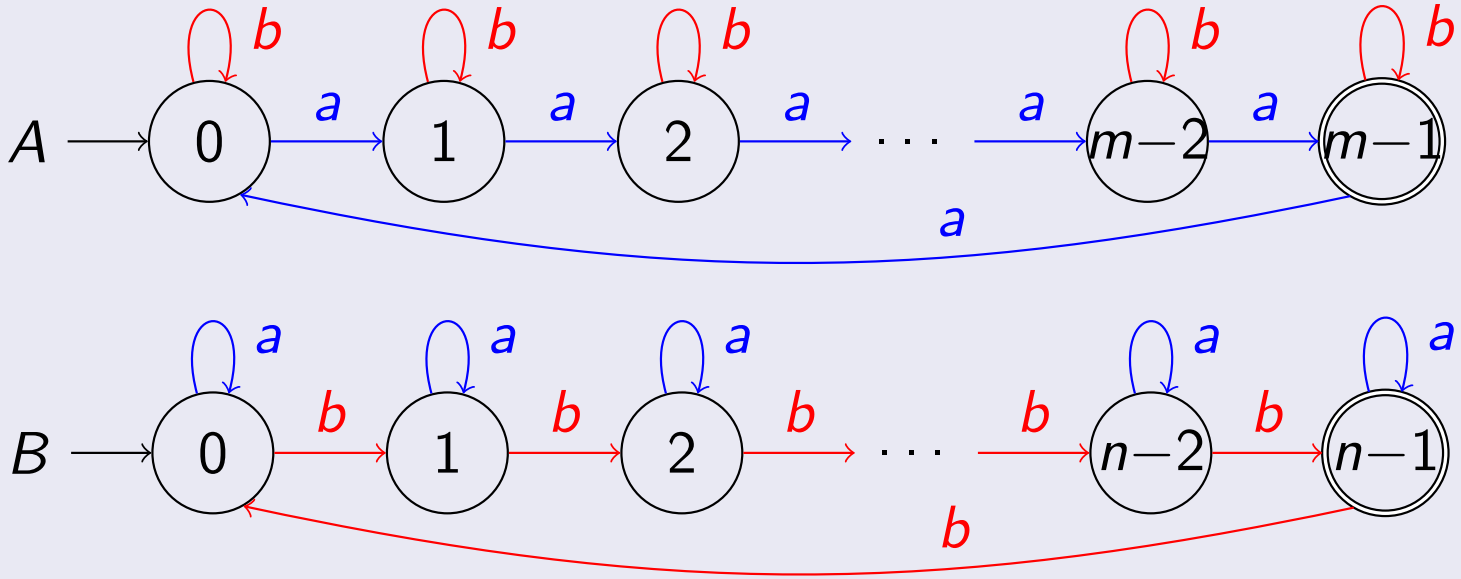
- We replace m equivalent states with a single state
- This decreases the upper bound to $m2^n - 2^{n-1} - m + 1$

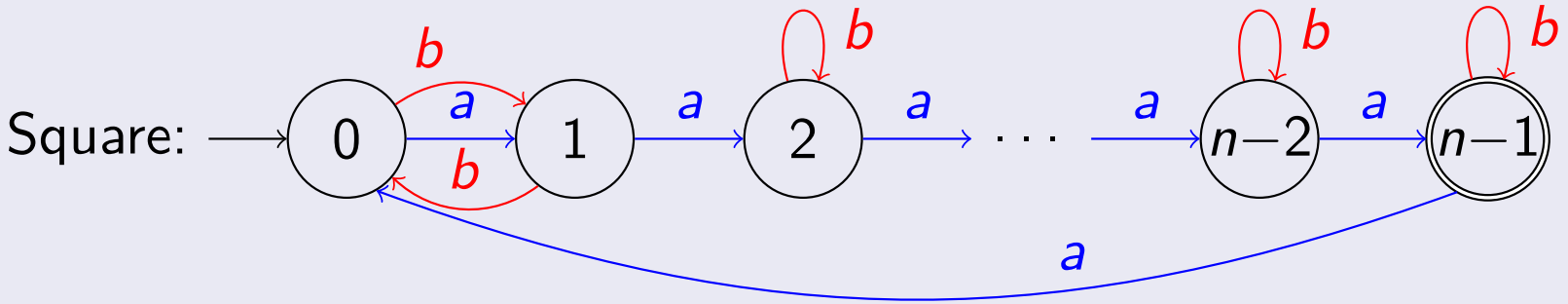
Square: $L(A)^2$
Regular upper bound: $n2^n - 2^{n-1}$

- For each reachable set $\{q\} \cup S$ we have
  - if $s \in F$, then $q \in S$;
  - if $s \notin F$, then $q \notin S$ (the contrary would be in contradiction with the permutation property of A)
- This makes half of usually reachable states in the DFA for $L(A)^2$ unreachable
- The new upper bound is $n2^{n-1} - 2^{n-2}$

Lower Bound Witnesses for Concatenation and Square

Concatenation (also witnesses for Boolean operations and cut):





The Cut Operation

Definition (Berglund, Björklund, Drewes, van d. Merwe, Watson 2013)

The cut of languages K and L over Σ is the language

$K \,!\, L = \{uv \mid u \in K, v \in L, uv' \notin K \text{ for every non-empty prefix } v' \text{ of } v\}$

Example: $abbc \in \{a, ab\}!\{bc\}$, but $abc \notin \{a, ab\}!\{bc\}$

The complexity of cut on DFAs (Drewes, Holzer, Jakobi, v.d.M. 2017)

$\lvert\Sigma\rvert \geq 2$	$(m-1)n+m$	with permutation DFA witnesses
$\lvert\Sigma\rvert = 1$	$\max\{2m-1, m+n-2\}$	

The complexity of cut on permutation DFAs

$\lvert\Sigma\rvert \geq 2$	$(m-1)n+m$
$\lvert\Sigma\rvert = 1$	$2m-1$ if there is ℓ in $\{2, 3, \ldots, n\}$ such that $m \bmod \ell \neq 0$;
	$2m-2$ otherwise

The Cut Operation

Definition (Berglund, Björklund, Drewes, van d. Merwe, Watson 2013)

The cut of languages K and L over Σ is the language

$K \,!\, L = \{uv \mid u \in K, v \in L, uv' \notin K \text{ for every non-empty prefix } v' \text{ of } v\}$

Example: $abbc \in \{a, ab\}!\{bc\}$, but $abc \notin \{a, ab\}!\{bc\}$

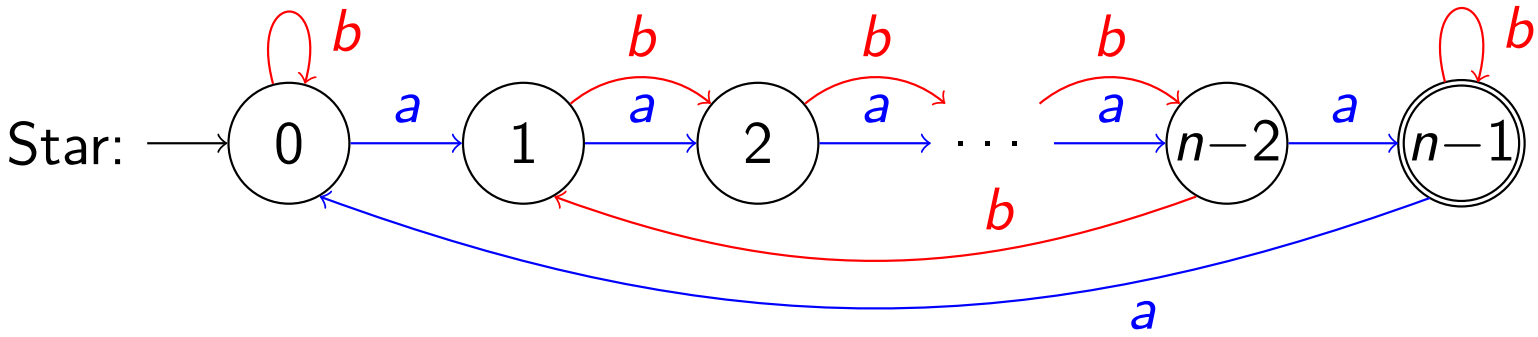
The complexity of cut on DFAs (Drewes, Holzer, Jakobi, v.d.M. 2017)

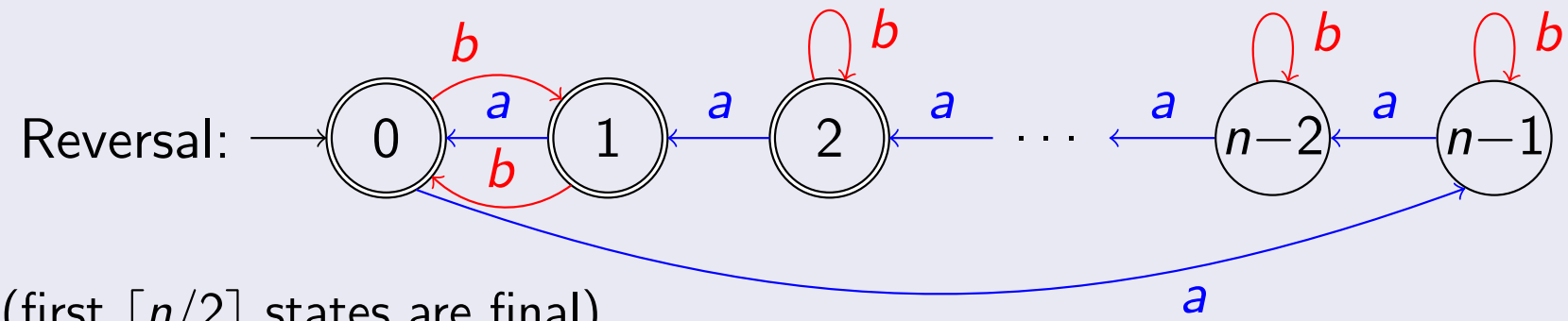
$\lvert\Sigma\rvert \geq 2$	$(m-1)n+m$	with permutation DFA witnesses
$\lvert\Sigma\rvert = 1$	$\max\{2m-1, m+n-2\}$	

The complexity of cut on permutation DFAs

$\lvert\Sigma\rvert \geq 2$	$(m-1)n+m$
$\lvert\Sigma\rvert = 1$	$2m-1$ if there is ℓ in $\{2,3,\ldots,n\}$ such that $m \bmod \ell \neq 0$;
	$2m-2$ otherwise

Lower Bound Witnesses for Other Operations





(first $\lceil n/2 \rceil$ states are final)

Left and right quotient:

- both witness DFAs have $k = \min\{m, n\}$ states and two symbols (a: cyclic permutation, b: transposition)
- modify the set of initial (left) or final (right) states in A
- the complexity is $\binom{k}{\lceil k/2 \rceil}$ for left, k for right, tight if $m \leq n$

Results: (Non-)Closure Properties and State Complexity in the Class of Group Languages

Operation	**Closed?**	**sc**	$\lvert\Sigma\rvert$	**sc,** $\lvert\Sigma\rvert = 1$
L^c	Yes [T68]	n	1	n
Boolean	Yes [T68]	mn	2	mn; $\gcd(m,n) = 1$
$K \,!\, L$	No	$(m-1)n+m$	2	$2m-1$ or $2m-2$
L^+	No	$\frac{3}{4}2^n - 1$	2	$(n-1)^2+1$
L^*	No [T68]	$\frac{3}{4}2^n$	2	$(n-1)^2+1$
KL	No [T68]	$m2^n - 2^{n-1} - m + 1$	2	mn; $\gcd(m,n) = 1$
L^2	No	$n2^{n-1} - 2^{n-2}$	2	$2n-1$
L^R	Yes [T68]	$\binom{n}{\lceil n/2 \rceil}$	2	n
$L^{-1}K$	Yes [T68]	$\binom{m}{\lceil m/2 \rceil}$; $m \leq n$	2	$\min(m,n)$
KL^{-1}	Yes [T68]	m; $m \leq n$	1	$\min(m,n)$
L^k	No	?		$k(n-1)+1$
$K \sqcup\!\sqcup L$	No	?		mn; $\gcd(m,n) = 1$
shift(L)	No	?		m
per(L)	No	?		m

Open Problems

1. What is the state complexity of quotients if $m > n$?
2. What is $\mathrm{sc}(L^k)$, $\mathrm{sc}(K \sqcup\!\sqcup L)$, and $\mathrm{sc}(\mathrm{shift}(L))$ if $|\Sigma| \geq 2$?
3. If L is a group language, then $\mathrm{per}(L)$ is always regular (Cano, Guaiana, Pin, 2013). So, what is $\mathrm{sc}(\mathrm{per}(L))$?

Partial answer to (3):

S. Hoffmann: State complexity bounds for the commutative closure of group languages. In: Proc. DCFS 2020, pp. 64-77

Theorem 4: Let $A = (Q, \{a_1, a_2, \ldots, a_k\}, \cdot, s, F)$.
Then an upper bound on $\mathrm{sc}(\mathrm{per}(L(A)))$ is

$$|Q|^k \left(\prod_{j=1}^{k} L_j\right)$$

where L_j is the least common multiple of lengths of cycles in the permutation of symbol a_j.

Thank you
for attention

Obrigado
pela atenção

Danke

Merci

ありがとう

감사합니다

Grazie

תודה

Ďakujem