Concatenation on Deterministic and Alternating Automata

Michal Hospodár, Galina Jirásková

Mathematical Institute
Slovak Academy of Sciences
Košice, Slovakia

NCMA 2016, Debrecen, Hungary

Outline

1. Concatenation on DFAs – upper bound
2. Concatenation on DFAs – more final states
3. Relation between DFAs and AFAs
4. DFA witnesses from literature revisited
5. New DFA witnesses with more final states
6. Concatenation on AFAs – tight bound

Concatenation

Definition

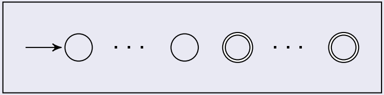
Concatenation of languages: $KL = \{uv \mid u \in K \text{ and } v \in L\}$

Regularity

If K and L are regular, then KL is regular.

Construction of a NFA for KL

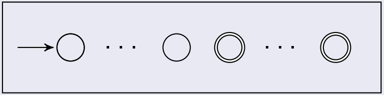
$A = (Q_A, \Sigma, \delta_A, s_A, F_A)$



$|Q_A| = m$

$|F_A| = k$

$B = (Q_B, \Sigma, \delta_B, s_B, F_B)$



$|Q_B| = n$

$|F_B| = \ell$

Concatenation

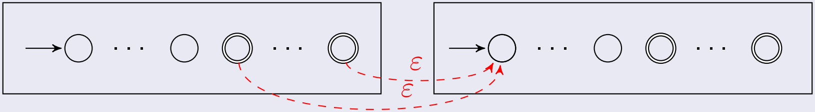
Definition

Concatenation of languages: $KL = \{uv \mid u \in K \text{ and } v \in L\}$

Regularity

If K and L are regular, then KL is regular.

Construction of a NFA for KL



$|Q_A| = m$
$|F_A| = k$

$|Q_B| = n$
$|F_B| = \ell$

Definition

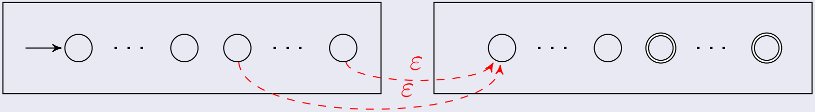
Concatenation of languages: $KL = \{uv \mid u \in K \text{ and } v \in L\}$

Regularity

If K and L are regular, then KL is regular.

Construction of a NFA for KL

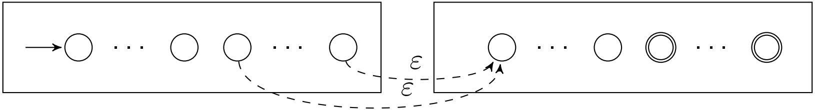
$$N = (Q_A \cup Q_B, \Sigma, \delta, s_A, F_B)$$



$|Q_A| = m$
$|F_A| = k$

$|Q_B| = n$
$|F_B| = \ell$

Concatenation of Regular Languages



In the subset automaton of NFA N

- every reachable subset is of the form $\{q\} \cup S$ where $q \in Q_A$ and $S \subseteq Q_B$
- if $q \in F_A$ and $s_B \notin S$, then $\{q\} \cup S$ is not reachable

Concatenation of Regular Languages

In the subset automaton of NFA N

- every reachable subset is of the form $\{q\} \cup S$ where $q \in Q_A$ and $S \subseteq Q_B$
- if $q \in F_A$ and $s_B \notin S$, then $\{q\} \cup S$ is not reachable

Complexity of concatenation:

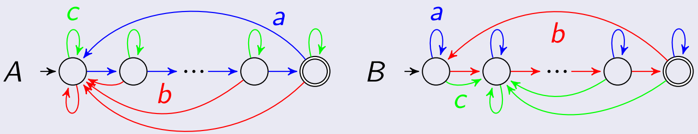
- $\leq m2^n - k2^{n-1}$ where $k = |F_A|$ [Yu, Zhuang, Salomaa 1994]
- depends on $k = |F_A|$, but does not depend on $\ell = |F_B|$

The value is maximal if $k = 1$: Upper Bound

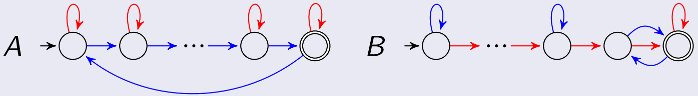
- $m2^n - 2^{n-1}$ [Maslov 1970]

Worst-Case Examples

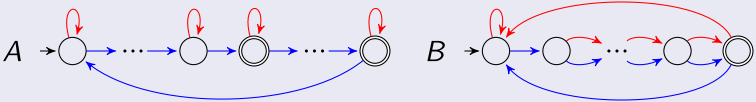
Ternary witnesses from Yu, Zhuang, Salomaa 1994



Binary witnesses from Maslov 1970



Binary witnesses for each k from Jirásek et al. 2005



Concatenation of Regular Languages

Complexity of concatenation:

- $\leq m2^n - k2^{n-1}$ where $k = |F_A|$ [Yu, Zhuang, Salomaa 1994]
- depends on $k = |F_A|$, but does not depend on $\ell = |F_B|$

In this paper:

- Are the bounds $m2^n - k2^{n-1}$ tight for all m, n, k, ℓ?
- Reachability does not depend on ℓ, but distinguishability does.

Why More Final States in DFA B?

Motivation

A. Fellah, H. Jürgensen, S. Yu: *Constructions for alternating finite automata*. Intern. J. Computer Math. 35 (1990), 117-132.

- *"...every n-state AFA has an equivalent NFA with at most $2^n + 1$ states. We conjecture that this bound is tight. However, we do not have any proof."*
- *"...we show that $2^m + n + 1$ states suffice for an AFA to accept the concatenation of two languages accepted by AFA with m and n states, respectively. We conjecture that this number of states is actually necessary in the worst case, but have no proof."*

Why More Final States in DFA B?

Motivation

A. Fellah, H. Jürgensen, S. Yu: *Constructions for alternating finite automata*. Intern. J. Computer Math. 35 (1990), 117-132.

- *"...every n-state AFA has an equivalent NFA with at most $2^n + 1$ states. We conjecture that this bound is tight. However, we do not have any proof."*
  – solved positively in Jirásková, CSR 2012
- *"...we show that $2^m + n + 1$ states suffice for an AFA to accept the concatenation of two languages accepted by AFA with m and n states, respectively. We conjecture that this number of states is actually necessary in the worst case, but have no proof."*
  – claimed to be solved in Jirásková, CSR 2012
  – the proof does not work

Alternating Finite Automata (AFAs)

$$A = (Q, \Sigma, \delta, s, F)$$

- Q is a non-empty finite set of states
- Σ is an input alphabet
- $s \in Q$ is the starting state
- $F \subseteq Q$ is the set of final states
- δ is the transition function that maps $Q \times \Sigma$ to
- a single state in DFA
- a union of states in NFA
- a boolean function of states in AFA

Alternating Finite Automata (AFAs)

Deterministic FA

δ	a	b
q_1	q_2	q_1
q_2	q_1	q_1

$\delta(q_1, aba) =$
$\delta(q_2, ba) =$
$\delta(q_1, a) = q_2$ – final state

Nondeterministic FA

δ	a	b
q_1	$q_1 \vee q_2$	F
q_2	q_2	q_1

$\delta(q_1, aba) =$
$\delta(q_1 \vee q_2, ba) =$
$\delta(q_1, a) = q_1 \vee q_2$ – contains final state

Alternating FA

δ	a	b
q_1	$q_1 \vee q_2$	q_1
q_2	$q_1 \wedge q_2$	$\overline{q_2}$

$\delta(q_1, aba) =$
$\delta(q_1 \vee q_2, ba) =$
$\delta(q_1 \vee \overline{q_2}, a) = q_1 \vee q_2 \vee \overline{q_1 \wedge q_2} = T$
– evaluate at $f = (0, 1)$
– gives 1 ... accepts *aba*

Relation between AFAs and DFAs

Lemma 5.1, 5.3 (cf. Fellah et al. 1990, Theorem 4.1)

Language L has an AFA with n states
iff L^R has a DFA with 2^n states, 2^{n-1} of them final.

Corollary 5.2

For every regular language L, we have $\mathrm{asc}(L) \geq \lceil \log(\mathrm{sc}(L^R)) \rceil$.

Jirásková: Descriptional Complexity of Operations on Alternating and Boolean Automata (CSR 2012)

claimed to solve the problem from Fellah et al. 1990
used binary languages from Jirásek, Jirásková, Szabari (JJS) 2005
the proof of JJS 2005 was incorrect

Results from Literature Revisited

Do the witnesses work for $\ell \geq 2$?

No. We gave $k = m/2$, $\ell = n/2$, and this is result:

bound	2	4	6
2	6	24	96
4	12	48	192
6	18	72	288

YZS	2	4	6
2	6	14	27
4	12	28	54
6	18	42	81

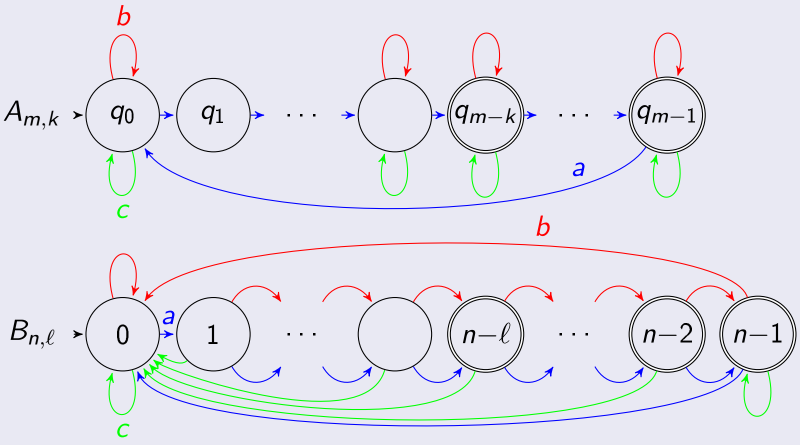
Maslov	2	4	6
2	5	4	18
4	10	5	35
6	15	6	52

JJS	2	4	6
2	6	22	84
4	12	42	156
6	18	63	225

Adding more final states in B destroys distinguishability.

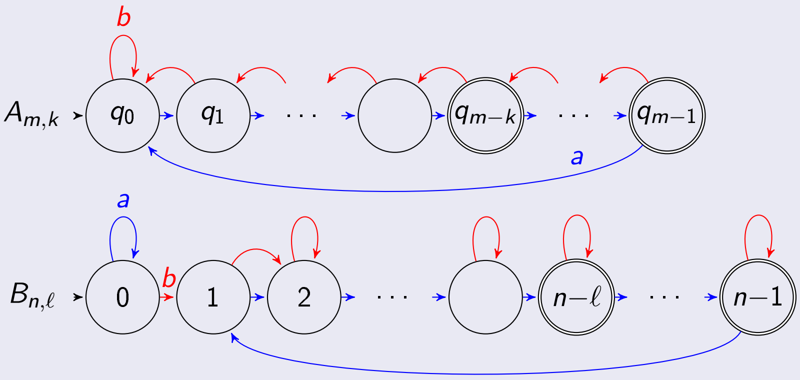
Ternary witness (based on JJS 2005)

Adding more final states in B does not affect reachability.
Distinguishability is secured by the letter c.



Binary witness (based on Dorčák 2015)

Limitations: $k \leq m - 2$ $m \geq 3$ $n \geq 4$



Worst-Case Example for AFA Concatenation

For every regular language L, we have $\operatorname{asc}(L) \geq \lceil \log(\operatorname{sc}(L^R)) \rceil$.

In order to have large AFA for KL,
we look for large DFA for $(KL)^R = L^R K^R$

Binary worst-case examples for DFA concatenation

L^R will be $A_{2^n,2^{n-1}}$
K^R will be $B_{2^m,2^{m-1}}$

Since L^R and K^R are worst cases,
$\operatorname{sc}(L^R K^R) = 2^n \cdot 2^{2^m} - 2^{n-1} \cdot 2^{2^m-1} = 2^{n-1} \cdot 2^{2^m}(1 + 1/2)$.

By Corollary 5.2,

$\operatorname{asc}(KL) \geq \lceil \log(2^{n-1} \cdot 2^{2^m}(1 + 1/2)) \rceil = 2^m + n$.

How many final states is in the DFA for $L^R K^R$?

In the minimal DFA for $L^R K^R$,

- a state $\{q\} \cup S$ is final iff $S \cap F_B \neq \emptyset$
- $S \cap F_B = \emptyset$ iff $S \subseteq (Q_B \setminus F_B)$
- $|Q_B| = 2^m$, $|F_B| = 2^{m-1} \Rightarrow |Q_B \setminus F_B| = 2^{m-1}$
- $|2^{Q_B \setminus F_B}| = 2^{2^{m-1}}$
- The minimal DFA for $(KL)^R$ has $2^{n-1}2^{2^m} + 2^{n-1}2^{2^m-1}$ states, of which $2^{n-1}2^{2^{m-1}} + 2^{n-1}2^{2^{m-1}-1}$ are non-final
- Since $m \geq 2$, we get $|F| \geq 2^{2^m+n-1}$.
- This is in contradiction with $\operatorname{asc}(KL) = 2^m + n$
- It follows that $\operatorname{asc}(KL) \geq 2^m + n + 1$,
  which solves the open problem from FJY 1990.

Summary, open problems

- Yu, Zhuang, Salomaa 1994 – proof for k final states
- Maslov 1970 – proof for k final states
- Jirásek, Jirásková, Szabari 2005 – new proof for k final states
- Ternary pair of languages – proof for k and ℓ final states
- Binary pair of languages – proof for k and ℓ final states with $k \leq m - 2$
- AFA concatenation – proof of upper bound from Fellah, Jürgensen, Yu 1990 (correction of Jirásková 2012)

Open problems

- AFA square binary witness
  (ternary: Čevorová, Jirásková, Krajňáková, CIAA 2014)
- DFA binary concatenation with $k = m - 1$

Thank You For Attention

Köszonjük a figyelmet

Paldies	Xièxiè
Teşekkür	Terima
Vielen Dank	Rahmat
Mulţumesc	Děkujeme
Obrigado	$E\upsilon\chi\alpha\rho\iota\sigma\tau\omega$
Ďakujeme	Dziękujemy
Kamsahabnida	Grazie