State complexity of complementing unambiguous finite automata

Michael Raskin, raskin@mccme.ru

LaBRI, Université de Bordeaux

July 10, 2018

Non-determinism in automata

The basic classes: deterministic and non-deterministic finite automata

The set of languages is the same

State complexity (number of states required) differs

Automata languages

$$L(\mathrm{DFA}) = L(\mathrm{NFA})$$

Non-determinism in automata: state complexity

Automata languages

$$L(\mathrm{DFA}) = L(\mathrm{NFA})$$

Exponentially more succinct

Non-determinism in automata: state complexity

Automata languages

$$L(\mathrm{DFA}) = L(\mathrm{NFA})$$

Exponentially more succinct

Intersection, union

Quadratic state complexity

Complement

No extra cost

Exponential state complexity

Reversing direction (left to right/right to left)

Exponential state complexity

No extra cost

Codeterministic automata

coDFA: DFA reading the word right-to-left

Union/intersection between DFA and coDFA —
exponential state complexity
(if we want to stay in DFA $\cup$ coDFA)

Unambiguous automata

UFA: NFA with at most one accepting run for each word

DFA	UFA	NFA
	more succinct	even more succinct
	Intersection, union	
$(\cdot)^2$	$(\cdot)^2$, ?	$(\cdot)^2$
	Complement	
No extra cost	?	$\mathit{exp}(\cdot)$
	Reversing order	
$\mathit{exp}(\cdot)$	No extra cost	No extra cost

Complementing UFA

- Known to be at least quadratic
- Lower bound holds for unary case
- Conjectured to be polynomial

Main result

- State complexity of complementing UFA:
  superpolynomial lower bound

Main result

- State complexity of complementing UFA:
  superpolynomial lower bound
- Even in unary case
- Even if complement is general NFA
- Even if language is also easy for DFA with multiple passes

Main result

- State complexity of complementing UFA:
  superpolynomial lower bound
- Even in unary case
- Even if complement is general NFA
- Even if language is also easy for DFA with multiple passes

Very weakly superpolynomial lower bound: $n^{\Omega(\log\log\log n)}$

Main result

- State complexity of complementing UFA:
  superpolynomial lower bound
- Even in unary case
- Even if complement is general NFA
- Even if language is also easy for DFA with multiple passes

Very weakly superpolynomial lower bound: $n^{\Omega(\log\log\log n)}$

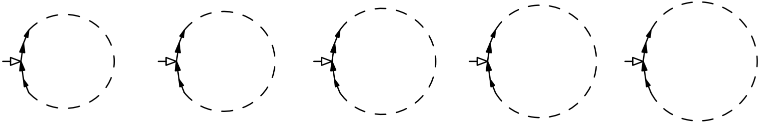
A lower bound for complementing a unary UFA **must** be weak:
upper bound $n^{O(\log n)}$
(Dębski, 2017)

Why?

Direct construction

Simple Chrobak normal form:
unary NFA := collection of cycles $C_1, \ldots, C_n$

Input word $\equiv$ length $\equiv$ remainder modulo $lcm(|C_1|, \ldots, |C_n|)$



Why? Tournaments!

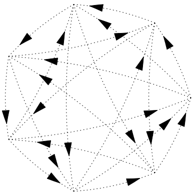
Input word $\equiv$ remainder modulo lcm of cycle lengths

Remainder 0 not in language; separation instead of complement
Square-free cycle lengths

Unambiguity:
Remainder 0 modulo $gcd(|C_i|, |C_j|)$ rules out acceptance by C_i or by C_j

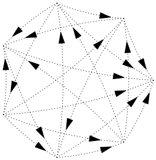
C_i yields to C_j:
remainder 0 modulo $gcd(|C_i|, |C_j|)$ rules out acceptance by C_i

Tournament of yielding between cycles

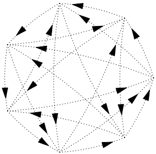


Why? Tournament properties

Bad case: Everyone yields to C_1, remainder 0 modulo C_1 separates



Good case: Every small set of cycles yields to some other cycle



Random tournament: good case

Why?

Input $\equiv$ remainder

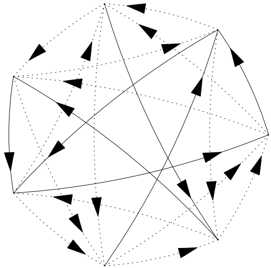
Tournament: Yielding between cycles

Random tournament is good

Technical details: tournament of yielding can be controlled

Lower bound for construction

Separation ≈ proof of non-inclusion ≈ bad remainder for some modulo
A proof of non-inclusion proves that every cycle yields



No small dominating set: many independent edges among the chosen

Lower bound for construction

A proof of non-inclusion proves that every cycle yields
No small dominating set: many independent edges among the chosen

Choice of accepting states: *gcd*, corresponding to a chosen edge, divides separating modulo

Careful assignment of prime factors

A lot of different *gcd*'s divide the length of a cycle ⇒ superpolynomial size

Future directions

- Non-unary case: is the state complexity exponential?
  Hypothesis: at least $2^{n^{\Theta(1)}}$
- Unary case: is the state complexity $n^{\Theta(\log n)}$?
- $\mathrm{DFA} \times \mathrm{coDFA}$: studied as transducers (bimachines) —
  how do automata behave?

Thanks for your attention.

Questions?