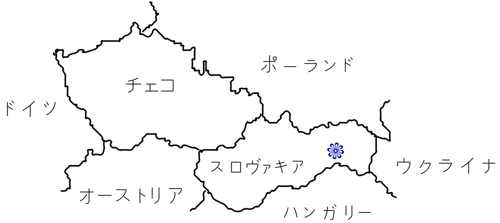
On the State Complexity of Complements, Stars, and Reversals of Regular Languages

Galina Jirásková

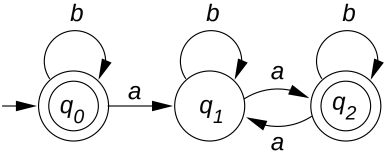
Mathematical Institute, Slovak Academy of Sciences
Košice, Slovakia



Outline

1. Basic Notions and Known Results
   - Deterministic and Nondeterministic Finite Automata
   - State Complexity
   - NFA to DFA Conversion: "Magic Numbers"
2. Main Results
   - Nondeterministic State Complexity of Complements
   - State Complexity of Stars and Reversals
   - Nondeterministic State Complexity of Stars and Reversals
3. Summary and Open Problems

Deterministic Finite Automata



A DFA $M = (Q, \Sigma, \delta, q_0, F)$

- $\delta : Q \times \Sigma \to Q$
- complete

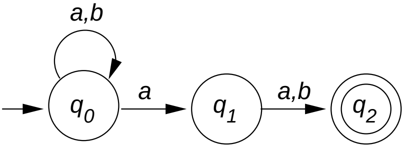
Definition

The state complexity of a regular language L, $\mathrm{sc}(L)$,
is the least number of states in any DFA accepting L.

Example

$L_{even} = \{w \in \{a, b\}^* \mid w \text{ has an even number of } a\text{'s}\}$
$\mathrm{sc}(L_{even}) = 2$

Nondeterministic Finite Automata



An NFA $N = (Q, \Sigma, \delta, q_0, F)$

- $\delta : Q \times \Sigma \to 2^Q$
- ε-free
- single initial state

Definition

The nondeterministic state complexity of a language L, $\text{nsc}(L)$, is the least number of states in any NFA accepting L.

Example

$L_2 = \{w \in \{a, b\}^* \mid w \text{ has an } a \text{ in the 2nd position from the end}\}$
$textnsc(L_2) = 3$
$\text{nsc}(L_k) = k + 1$ and $sc(L_k) = 2^k$

NFA to DFA Conversion

Theorem (Rabin, Scott 1959)

Every NFA has an equivalent DFA.

- Subset construction: n-state NFA $\rightarrow$ at most 2^n-state DFA
- Tight in binary case [Lupanov 1963, Moore 1971, ...]
- Unary case: at most $F(n) \approx e^{\sqrt{n \ln n}}$ states [Chrobak 1986]

Example

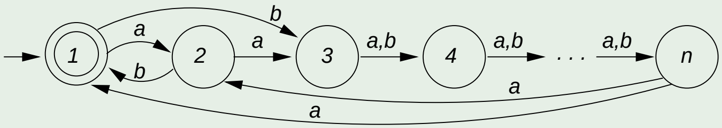


Figure: A binary NFA which needs 2^n deterministic states

NFA to DFA: "Magic Numbers"

Iwama et al., DLT 1997 and MFCS 2000:

- Is there an n-state NFA whose equivalent minimal DFA has α states for each α with $n \leqslant \alpha \leqslant 2^n$?
- Possible holes are called "magic numbers".
- The numbers $2^n - 2^k$, $2^n - 2^k - 1$ ($k \leqslant n/2$), and $2^n - \ell$ ($\ell \leqslant 2n$, coprimality) are not magic.

NFA to DFA: "Magic Numbers"

Question

Is there an n-state NFA whose equivalent minimal DFA has exactly α states for each α with $n \leqslant \alpha \leqslant 2^n$?

- Yes, alphabet of size 2^n [Jirásková, MFCS 2001]
- Yes, alphabet of size $n + 2$ [Geffert, DCFS 2005]
- Yes, four-letter alphabet [Jirásek et al., DLT 2007]
- No, for a unary alphabet:
  - All numbers from $F(n)$ to 2^n are magic [Chrobak 1986]
  - A lot of magic numbers from n to $F(n)$ [Geffert, MFCS 2006]

Complements: Nondeterministic State Complexity

A Similar Question on Complements

Is there an n-state NFA language L
such that every minimal NFA for the language L^c
has exactly α states for each α with $\log n \leqslant \alpha \leqslant 2^n$?

- Yes, alphabet of size 2^{n+1} [Jirásek et al., CIAA 2004]
- Yes, alphabet of size $2n$ [Szabari, ITAT 2006]

Complements: Five-Letter Case

Theorem

For all n and α with $\log n \leqslant \alpha \leqslant 2^n$, there exists a regular language L over a five-letter alphabet such that $\text{nsc}(L) = n$ and $\text{nsc}(L^c) = \alpha$.

Complements: Five-Letter Case

Theorem

For all n and α with $\log n \leqslant \alpha \leqslant 2^n$,
there exists a regular language L over a five-letter alphabet
such that $\mathrm{nsc}(L) = n$ and $\mathrm{nsc}(L^c) = \alpha$.

Sketch of proof.

- Take an n-state NFA with a four-letter alphabet whose equivalent minimal DFA has α states [DLT 2007].
- Add transitions on a new symbol so that the reachable states are the same. Let L be the language accepted by this NFA.
- For each reachable set S, define a pair (x_S, y_S) so that
  (1) $x_S y_S \in L^c$ for each S;
  (2) if $S \neq T$, then $x_S y_T \notin L^c$ or $x_T y_S \notin L^c$. □

Complements: Example

Example

The case of $n = 6$ and $\alpha = 24$:

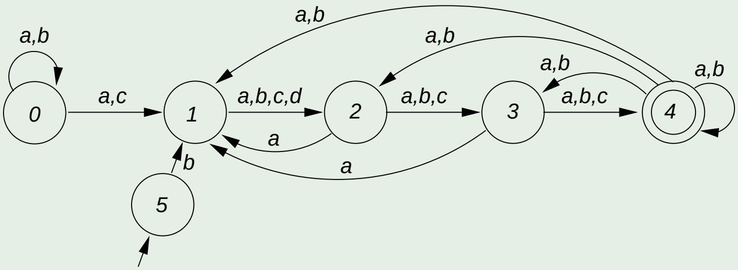


Figure: A 6-state NFA which needs 24 deterministic states [DLT 2007].

Complements: Example

Example

The case of $n = 6$ and $\alpha = 24$:

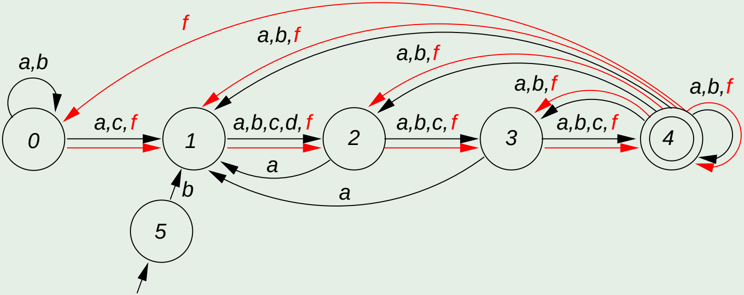


Figure: A 6-state NFA M over $\{a, b, c, d, f\}$ with $\text{nsc}(L(M)^c) = 24$.

Complements: Binary Case

Theorem

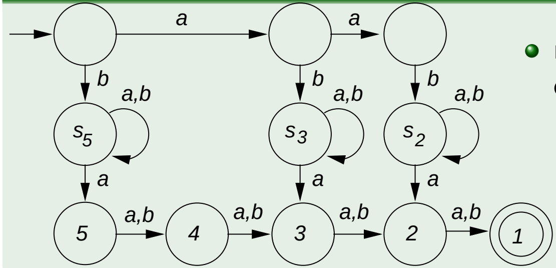
For all n and α with $3 \log n \leqslant \alpha \leqslant 2^{n/3}$, there exists a binary regular language L such that $\text{nsc}(L) = n$ and $\text{nsc}(L^c) = \alpha$.

Complements: Binary Case

Theorem

For all n and α with $3 \log n \leqslant \alpha \leqslant 2^{n/3}$, there exists a binary regular language L such that $\text{nsc}(L) = n$ and $\text{nsc}(L^c) = \alpha$.

Example



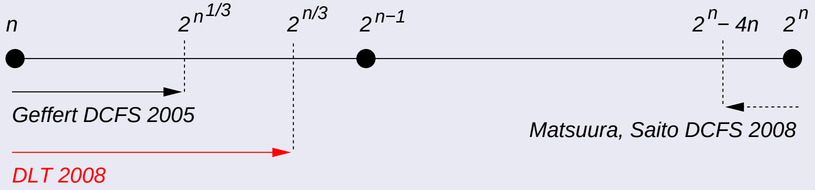
- needs $3 + 2^5 + 2^3 + 2^2$ deterministic states

Magic Numbers: Binary Case

Corollary

All numbers from n to $2^{n/3}$ are non-magic for NFA to DFA conversion in a binary case.

Non-Magic Numbers in Binary Case



Stars: State Complexity

Stars: State Complexity

Theorem

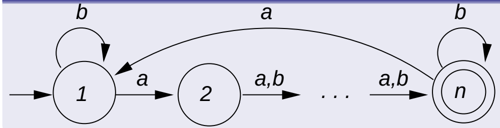
*For all n and α with $1 \leqslant \alpha \leqslant \frac{3}{4} \cdot 2^n$,
there exists a regular language (over an exponential alphabet)
such that $\mathrm{sc}(L) = n$ and $\mathrm{sc}(L^*) = \alpha$.*

Stars: State Complexity

Theorem

For all n and α *with* $1 \leqslant \alpha \leqslant \frac{3}{4} \cdot 2^n$*,*
there exists a regular language (over an exponential alphabet)
such that $\mathrm{sc}(L) = n$ *and* $\mathrm{sc}(L^*) = \alpha$*.*

Proof idea



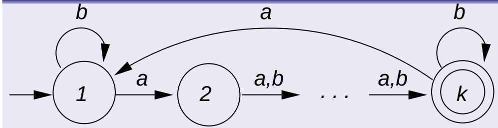
- A binary DFA for $\alpha = \frac{3}{4} \cdot 2^n$ [Yu et al. 1984]

Stars: State Complexity

Theorem

For all n and α *with* $1 \leqslant \alpha \leqslant \frac{3}{4} \cdot 2^n$*,
there exists a regular language (over an exponential alphabet)
such that* $\mathrm{sc}(L) = n$ *and* $\mathrm{sc}(L^*) = \alpha$*.*

Proof idea



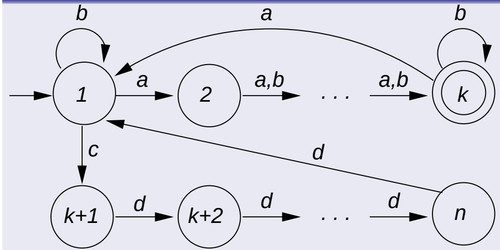
- The worst case example on k states
-
-

Stars: State Complexity

Theorem

For all n and α with $1 \leqslant \alpha \leqslant \frac{3}{4} \cdot 2^n$, there exists a regular language (over an exponential alphabet) such that $\mathrm{sc}(L) = n$ and $\mathrm{sc}(L^*) = \alpha$.

Proof idea



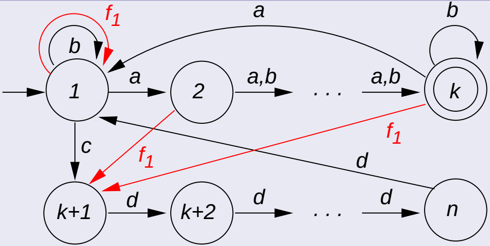
- The worst case DFA on k states
- Two more symbols for $\alpha = n - k + \frac{3}{4} \cdot 2^k$
-

Stars: State Complexity

Theorem

For all n and α *with* $1 \leqslant \alpha \leqslant \frac{3}{4} \cdot 2^n$, *there exists a regular language (over an exponential alphabet) such that* $\mathrm{sc}(L) = n$ *and* $\mathrm{sc}(L^*) = \alpha$.

Proof idea



- The worst case DFA on k states
- Two more symbols for $\alpha = n - k + \frac{3}{4} \cdot 2^k$
- Add m more symbols for $\alpha = n - k + \frac{3}{4} \cdot 2^k + m$

ytixelpmoC etatS :slasreveR

Reversals: State Complexity

Reversals: State Complexity

Theorem

For all n *and* α *with* $\log n \leqslant \alpha \leqslant 2^n$,
there exists a regular language (over an exponential alphabet)
such that $\mathrm{sc}(L) = n$ *and* $\mathrm{sc}(L^R) = \alpha$.

Reversals: State Complexity

Theorem

For all n and α with $\log n \leqslant \alpha \leqslant 2^n$, there exists a regular language (over an exponential alphabet) such that $\mathrm{sc}(L) = n$ and $\mathrm{sc}(L^R) = \alpha$.

Example

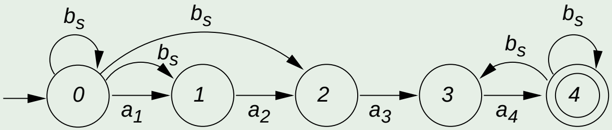


Figure: An NFA for the reversal; $S = \{0, 1, 2\}$

Stars and Reversals: Nondeterministic State Complexity

Theorem

For all n and α with $1 \leqslant \alpha \leqslant n+1$, there exists a binary regular language such that $\text{nsc}(L) = n$ and $\text{nsc}(L^*) = \alpha$.

Theorem

For all n and α with $\alpha \in \{n-1, n, n+1\}$, there exists a binary regular language such that $\text{nsc}(L) = n$ and $\text{nsc}(L^R) = \alpha$.

Summary and Open Problems

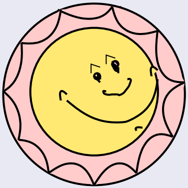
Our Results

	sc	Alphabet	nsc	Alphabet
L^c	$\{n\}$	arbitrary	$\log n$ .. 2^n	5
L^*	1 .. $\frac{3}{4} \cdot 2^n$	2^n	1 .. $n+1$	2
L^R	$\log n$ .. 2^n	2^n	$\{n-1, n, n+1\}$	2

Open Problems

- Complements over smaller alphabets (unary case)
- Stars and reversals over a fixed alphabet

Thank You for Your Attention



ありがとうございます