Nondeterministic state complexity of square root

S. Onishchenko*

We investigate the nondeterministic state complexity of the square-root operation $\sqrt{L} = \{ w \mid ww \in L \}$ on regular languages represented by nondeterministic finite automata. For an n-state NFA accepting L, it was previously known that $\sqrt{L}$ can be accepted by an NFA with at most n^3 states, while the best lower bound, was only $(n-1)(n-2)(n-3)$. We are going to close this gap completely and prove that n^3 states are sufficient and necessary in the worst case.

Keywords: state complexity, language operations, nondeterministic finite automata.

1. Preliminaries

Let Σ be a finite alphabet. We write Σ^* for the free monoid generated by Σ, and ε for the empty word.

A *nondeterministic finite automaton* (NFA) is a quintuple $A = (Q, \Sigma, \delta, Q_0, F)$, where Q is a finite set of states, Σ is the input alphabet, $Q_0 \subset Q$ is the set of initial states, $F \subseteq Q$ is the set of accepting states, and $\delta\colon Q \times \Sigma \to 2^Q$ is the transition function.

The transition function is extended in the standard way to a mapping $\delta\colon Q \times \Sigma^* \to 2^Q$ by setting $\delta(q, \varepsilon) = \{q\}$ for every $q \in Q$, and $\delta(q, wa) = \bigcup_{p \in \delta(q,w)} \delta(p, a)$ for every $q \in Q$, $w \in \Sigma^*$, and $a \in \Sigma$.

The language recognized by A is

$$L(A) = \{ w \in \Sigma^* \mid \exists q_0 \in Q_0\colon \delta(q_0, w) \cap F \neq \varnothing \}.$$

We shall use the fooling set technique to derive lower bounds on the nondeterministic state complexity.

Definition 1. Let $L \subseteq \Sigma^*$. A set

$$S = \{(x_i, y_i) \mid 1 \leq i \leq m\} \subseteq \Sigma^* \times \Sigma^*$$

is called a *fooling set* for L if the following conditions are satisfied:

1. $x_i y_i \in L$ for every $1 \leq i \leq m$;

* *Onishchenko Sergey* — Saint Petersburg Leonhard Euler International Mathematical Institute, e-mail: dimer03@mail.ru

Published under CC-BY 4.0

2. for every pair of distinct indices i, j, at least one of the words $x_i y_j$ and $x_j y_i$ does not belong to L.

The following theorem was proved by Jean-Camille Birget [1].

Theorem 1. *Let $L \subseteq \Sigma^*$. If L admits a fooling set of cardinality m, then every NFA recognizing L has at least m states.*

2. The upper bound

For the sake of completeness, here we will present that well-known construction. The idea of our lower-bound construction will be based on it.

Theorem 2. *For every NFA $A = (Q, \Sigma, \delta, Q_0, F)$ with n states, which recognizes language L, there exists an NFA B with n^3 states, which recognizes language $\sqrt{L}$.*

Full proof is represented in [2].

3. The lower bound

Theorem 3. *For every $n \geq 6$ there exists a NFA A_n with n states, which recognizes a language L_n such that $\sqrt{L_n}$ is not recognized by any NFA with fewer than n^3 states.*

References

[1] J.-C. Birget, "Intersection and union of regular languages and state complexity", *Information Processing Letters*, **43**:4 (1992), 185–190. DOI: 10.1016/0020-0190(92)90198-5.

[2] S. Onishchenko, "Nondeterministic state complexity of square root", 2026, arXiv: 2605.02957.

Received on September 11, 2026