# Об одной разновидности функции Ландау, возникающей при изучении однозначных конечных автоматов
# On a variant of Landau's function arising in the study of unambiguous finite automata*

Александр Охотин

30 декабря 2020 г.

**Краткое содержание.** Конечный автомат — это модель вычислений с конечной памятью: он читает поданную на вход строку слева направо, изменяя своё состояние при прочтении каждого очередного символа; состояние же, полученное после прочтения всей строки, определяет, принимается она или отвергается. В *детерминированном* автомате (DFA) состояние на каждом следующем шаге функционально зависит от текущего состояния и от читаемого символа; в *недетерминированном* (NFA) для данного текущего состояния и символа может быть задано несколько возможных состояний на следующем шаге, и строка считается принятой, если она принимается при какой-то последовательности выборов. Классический результат теории автоматов — преобразование всякого NFA с $n$ состояниями к DFA с $2^n$ состояниями (Рабин и Скотт [1959]), причём $2^n$ состояний в худшем случае необходимы (Лупанов [1963]). Если же входная строка задана над односимвольным алфавитом, то число состояний, необходимых для детерминизации, составляет $g(n) + O(n^2)$ (Любич [1964]; Хробак [1986]), где $g(n) = \max_{n=n_1+\ldots+n_k} \operatorname{lcm}(n_1, \ldots, n_k) = e^{(1+o(1))\sqrt{n \ln n}}$ — *функция Ландау* (Ландау [1903]).

В докладе будет рассказано о сложности детерминизации для одного подкласса недетерминированных автоматов — *однозначных* конечных автоматов (unambiguous; UFA) — таких, что если автомат принимает строку, то есть лишь единственная последовательность недетерминированных выборов, ведущая к успеху, в то время как отвергающих вычислений может быть сколь угодно много. Однозначный недетерминизм — это важное понятие теории сложности вычислений, и в простом случае конечных автоматов оно лучше поддаётся изучению. Для многосимвольных алфавитов сложность детерминизации не становится меньше — в худшем случае нужны все $2^n$ состояний (Люнг [2005]). Интересен случай односимвольного алфавита, и первый результат по этой задаче — нижняя оценка вида $e^{\Omega((\log n)^{3/2})}$ (Равикумар и Ибарра [1989]). Впоследствии оказалось, что для UFA с $n$ состояниями над односимвольным алфавитом, необходимое и достаточное число состояний в равносильных им DFA имеет вид $\max_\ell \widetilde{g}(n-\ell) + \ell$, где $\widetilde{g}$ — новая разновидность функции Ландау, для которой $\widetilde{g}(n)$ определяется как максимум $\operatorname{lcm}(p_1 + \ldots + p_k)$ по всем $p_1, \ldots, p_k$, где $p_1 + \ldots + p_k \leqslant n$, и существуют остатки $f_1, \ldots, f_k$, где $0 \leqslant f_i < p_i$, удовлетворяющие условию $f_i \not\equiv f_j \pmod{\gcd(p_i, p_j)}$ для всех $i \neq j$. Эта функция (OEIS A174234) растёт медленнее, чем классическая функция Ландау: $\widetilde{g}(n) = e^{\Theta(\sqrt[3]{n(\ln n)^2})}$ (Охотин [2012]). Вопрос о более точной оценке её асимптотики остаётся открытым.

*Доклад, прочитанный на новогоднем коллоквиуме факультета МКН СПбГУ.

В заключение будет рассказано о любопытной родственной задаче — преобразовании UFA над односимвольным алфавитом к UFA, распознающему дополнение — то есть принимающему ровно те строки, которые не принимает исходный автомат. Чтобы представить дополнение UFA с $n$ состояниями, заведомо достаточно $\max_\ell \widetilde{g}(n-\ell)+\ell$ состояний — но получить нижнюю оценку оказывается весьма нелегко. Первая полученная нижняя оценка — $n^{2-o(1)}$ (Охотин [2012]), впоследствии она была улучшена до $n^{(\log\log\log n)^{\Omega(1)}}$ (Раскин [2018]).

**Abstract.** A finite automaton is a model of computation with finite memory: it reads a string given as an input from left to right, changing its state upon reading every next symbol; then the state entered after reading the entire string determines whether the string is accepted or rejected. In a *deterministic* automaton (DFA), the state at every next step functionally depends on the previous state and on the symbol being read; in a *deterministic* automaton (NFA), for a current state and a symbol being read, more than one possible next state can be specified, and then the string is considered accepted if it is accepted for some sequence of choices. A classical result of automata theory is a transformation of every $n$-state NFA to a DFA with $2^n$ states (Rabin and Scott [1959]), while $2^n$ states are necessary in the worst case (Lupanov [1963]). If the input string is defined over a one-symbol alphabet, then the number of states required by the determinization procedure is $g(n)+O(n^2)$ (Lyubich [1964]; Chrobak [1986]), where $g(n) = \max_{n=n_1+\ldots+n_k} \operatorname{lcm}(n_1,\ldots,n_k) = e^{(1+o(1))\sqrt{n\ln n}}$ is *Landau's function* (Landau [1903]).

The subject of this talk is the determinization complexity for a certain subclass of nondeterministic automata: the *unambiguous* finite automata (UFA), defined by the condition that if an automaton accepts a string, then there is a unique sequence of nondeterministic choices leading to acceptance, whereas the number of rejecting computations is not restricted. Unambiguous nondeterminism is an important notion of the computational complexity theory, and it is more amenable to study in the simple case of finite automata. For multiple-symbol alphabet the determinization complexity is not any smaller: in the worst case, all $2^n$ states are still needed (Leung [2005]). The case of a one-symbol alphabet is non-trivial, and the first result on this problem was a lower bound of the form $e^{\Omega((\log n)^{3/2})}$ (Ravikumar and Ibarra [1989]). Later it turned out that for an $n$-state UFA over a unary alphabet, the necessary and sufficient number of states in the equivalent DFA is of the form $\max_\ell \widetilde{g}(n-\ell)+\ell$, where $\widetilde{g}$ is a new variant of Landau's function, with $\widetilde{g}(n)$ defined as the maximum of $\operatorname{lcm}(p_1+\ldots+p_k)$ over all $p_1,\ldots,p_k$ with $p_1+\ldots+p_k \leqslant n$, for which there exist residues $f_1,\ldots,f_k$, with $0 \leqslant f_i < p_i$, satisfying the condition $f_i \not\equiv f_j \pmod{\gcd(p_i,p_j)}$ for all $i \neq j$. This function (OEIS A174234) grows slower than the classical Landau's function: $\widetilde{g}(n) = e^{\Theta(\sqrt[3]{n(\ln n)^2})}$ (Okhotin [2012]). The question on a more precise asymptotical estimation of this function remains open.

In conclusion, an interesting related problem shall be discussed: the transformation of a UFA over a one-symbol alphabet to a UFA recognizing its complement, that is, accepting exactly those strings not accepted by the original automaton. In order to represent the complement of an $n$-state UFA, it is certainly sufficient to use $\max_\ell \widetilde{g}(n-\ell)+\ell$ states, but obtaining any lower bounds turns out to be difficult. The first lower bound obtained was $n^{2-o(1)}$ (Okhotin [2012]), and later it was improved to $n^{(\log\log\log n)^{\Omega(1)}}$ (Raskin [2018]).

## Список литературы

[1986] M. Chrobak, “Finite automata and unary languages”, *Theoretical Computer Science*, 47 (1986), 149–158.

[1903] E. Landau, “Über die Maximalordnung der Permutationen gegebenen Grades” (On the maximal order of permutations of a given degree), *Archiv der Mathematik und Physik, Ser. 3*, 5 (1903), 92–103.

[2005] H. Leung, "Descriptional complexity of NFA of different ambiguity", *International Journal of Foundations of Computer Science*, 16:5 (2005), 975–984.

[1963] O. B. Lupanov, "A comparison of two types of finite automata", in Russian, *Problemy Kibernetiki*, 9 (1963), 321–326.

[1964] Yu. Lyubich, "Bounds for the optimal determinization of nondeterministic autonomic automata", *Sibirskii Matematicheskii Zhurnal*, 2 (1964), 337–355, in Russian.

[2012] A. Okhotin, "Unambiguous finite automata over a unary alphabet", *Information and Computation*, 212 (2012), 15–36.

[1959] M. O. Rabin, D. Scott, "Finite automata and their decision problems", *IBM Journal of Research and Development*, 3:2 (1959), 114–125.

[2018] M. Raskin, "A superpolynomial lower bound for the size of non-deterministic complement of an unambiguous automaton", *45th International Colloquium on Automata, Languages, and Programming* (ICALP 2018, Prague, Czech Republic, July 9–13, 2018), LIPIcs 107.

[1989] B. Ravikumar, O. H. Ibarra, "Relating the type of ambiguity of finite automata to the succinctness of their representation", *SIAM Journal on Computing*, 18:6 (1989), 1263–1282.