О сложности перехода к правильному линейному виду

П. С. Дергач[1] Д. А. Сальцова[1]

Данная работа посвящена изучению правильного линейного вида для регулярных языков с полиномиальной функцией роста и улучшению соответствующих оценок на сложность перехода от линейного вида к правильному линейному виду.

Удалось понизить ранее известную из работы [1] оценку n^2 до оценки $\frac{n^2}{2}+n$. Так же получена верхняя оценка $\frac{n^2}{4}$ для языков вида $\beta_1^*\beta_2^*\dots\beta_s^*$, в которых слова β_i соизмеримы.

Ключевые слова: регулярный язык, функция роста, правильный линейный вид, сложность.

1. Введение

Известно, что регулярные языки с полиномиальной функцией роста задаются регулярными выражениями линейного вида [1]. Выражениями линейного вида называются конечные объединения выражений вида

$$P = \alpha_1 \cdot (\beta_1)^* \cdot \alpha_2 \dots \alpha_s \cdot (\beta_s)^* \cdot \alpha_{s+1},$$

где $s \in \mathbb{N}_0,\ \alpha_1, \dots, \alpha_{s+1} \in A^*,\ \beta_1, \dots, \beta_s \in A^* \setminus \{\lambda\}$.

Выражениями правильного линейного вида называются выражения этого же вида, для которых $\alpha_2, \dots, \alpha_s$ не пусты и первые буквы слов $\beta_i,\ \alpha_{i+1}$ различны при всех $1 \le i \le s$ (для пустого α_{s+1} это условие при $i = s$ опускается).

Необходимо оценить сложность перехода от представлений языков выражениями линейного вида к представлениям выражениями правильного линейного вида.

В статье мы сначала даем определения и основные свойства этих выражений, а также формулируем задачу об оценке выражений.

Ранее в работе [1] изучался этот вопрос и была получена квадратичная верхняя оценка. Также в работе [2] для случая несоизмеримых

[1] *Дергач Пётр Сергеевич* — к.ф.-м.н., м.н.с. каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: dergachpes@gmail.com.

Dergach Peter Sergeevich — Ph.D., junior researcher, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems.

[1] *Сальцова Диана Александровна* — выпускник Филиала МГУ имени М. В. Ломоносова в городе Ташкенте, e-mail: saltsovadiana@gmail.com.

Saltsova Diana Alexandrovna — Graduate of the M. V. Lomonosov Moscow State University Branch in Tashkent.

выражений линейного вида (частный случай) была получена линейная верхняя оценка $4n$.

Далее в статье мы улучшаем общую верхнюю оценку на сложность перехода от регулярного выражения линейного вида к регулярному выражению правильного линейного вида.

Также отдельно рассмотрен случай, когда язык имеет соизмеримый вид. Получена верхняя оценка $\frac{n^2}{4}$.

В заключении мы подводим итоги и обсуждаем возможные направления для будущих исследований в данной области.

2. Основные определения и результаты

Большая часть определений данного раздела взята из источников [3]-[5].

Определение 1. Дадим определение регулярного множества. Множество P, называется *регулярным* в алфавите А, если его можно получить из пустого множества и одноэлементных однобуквенных множеств $\{a\}$, $a \in A$ следующим образом

$\emptyset$ — регулярное множество в алфавите A;

$\{a\}$ — регулярное множество в алфавите A, a — произвольная буква алфавита A;

Если $P_1,\ P_2$ — регулярные множества в алфавите A, то и множества $P_1 \cup P_2,\ P_1 \cdot P_2,\ (P_1)^*$ будут регулярными множествами в алфавите A.

Определение 2. Пусть $P \subseteq A^*$ и $n \in \mathbb{N}$. Через $P_{\leq}(n)$ обозначаем множество всех слов из P длины не больше n. Через T_P обозначаем функцию $T_P : N \to N_0$, где

$$T_P(n) := |P_{\leq}(n)|$$

для всех $n \in \mathbb{N}$. Называем T_P *функцией роста* для P. Говорим, что P имеет *полиномиальную функцию роста*, если $T_P(n)$ ограничена сверху полиномом от n.

Определение 3. *Регулярным выражением* в алфавите A, называется слово из алфавита $A \cup \{^*,\ \cdot,\ \vee,\ (,),\ \lambda\}$, которое определяется следующим образом:

λ — регулярное выражение в алфавите A;

Буквы алфавита A — регулярные выражения в алфавите A;

Если $P_1,\ P_2$ — регулярные выражения в алфавите A, то выражения $P_1 \cup P_2,\ P_1 \cdot P_2,\ (P_1)^*$ — будут регулярными выражениями в алфавите A.

Определение 4. *Длиной регулярного выражения* P в алфавите A будем называть количество букв из алфавита A, которые входят в это регулярное выражение с учетом кратности.

Определение 5. Слово $\alpha \in A^*$ называется *измельчением* слова $\beta \in A^*$, если при некотором $k \in \mathbb{N}$ выполняется равенство $\beta = \alpha^k$.

Определение 6. Слово $\alpha \in A^*$ называется *минимальным измельчением* непустого слова $\beta \in A^*$, если его длина минимальна среди всех измельчений слова β.

Определение 7. Два слова $\beta_1,\ \beta_2 \in A^*$ называются *соизмеримыми*, если их минимальные измельчения совпадают. В противном случае будем называть их *несоизмеримыми*.

Определение 8. Семейство слов $\beta_1, \beta_2, \ldots, \beta_n \in A^*$ называется *соизмеримым в совокупности*, если у них есть общее измельчение.

Определение 9. Регулярное выражение P в алфавите A имеет *линейный вид*, если оно представимо в виде конечного объединения выражений вида

$$P = \alpha_1 \cdot (\beta_1)^* \cdot \alpha_2 \ldots \alpha_s \cdot (\beta_s)^* \cdot \alpha_{s+1},$$

где $s \in \mathbb{N}_0,\ \alpha_1, \ldots, \alpha_{s+1} \in A^*,\ \beta_1, \ldots, \beta_s \in A^* \setminus \{\lambda\}$.

Если при этом $\alpha_2, \ldots, \alpha_s$ не пусты и первые буквы слов $\beta_i,\ \alpha_{i+1}$ при всех $1 \leq i \leq s$ различны (для пустого α_{s+1} это условие при $i = s$ опускается), то будем называть такое выражение *выражением правильного линейного вида.*

Определение 10. Будем говорить, что выражение линейного вида *несоизмеримо*, если все соседние пары β_i, β_{i+1} в нём несоизмеримы.

Определение 11. Пусть регулярное множество P задано конечным объединением регулярных выражений линейного вида. *Линейной сложностью такого представления* называем максимальную из длин этих выражений.

Определение 12. Пусть регулярное множество P задано конечным объединением регулярных выражений правильного линейного вида. *Правильной линейной сложностью такого представления* называем максимальную из длин этих выражений.

Определение 13. Пусть $P \subseteq \mathbb{N}$. Называем такое множество периодическим, если существуют $T_0,\ T \in \mathbb{N}$ такие, что начиная с $k \geq T_0$ из условия $k \in P$ следует условие $k + T \in P$. Число T_0 называем *предпериодом*, а число T — *периодом.*

Определение 14. Пусть $P \subseteq \mathbb{N}$ — периодическое множество. Через $T_{\text{предп}}(P)$ будем обозначать его минимальный предпериод.

Определение 15. Пусть $A = (\nu^{a_1})^*(\nu^{a_2})^* \ldots (\nu^{a_s})^*$. Тогда через $X_\nu(A)$ обозначаем множество $X_\nu(A) = \{l \mid \nu^l \in A\}$.

Определение 16. Пусть $a, b \in \mathbb{N}$. Если a делится нацело на b, то пишем $b|a$. Через $Z_{a,b}$ обозначаем множество:

$$Z_{a,b} := \{n \in \mathbb{N} |\ n \geq a,\ b|n\}.$$

Теорема 1. *Пусть регулярный язык задаётся регулярным выражением следующего линейного вида*

$$A = (\nu^{a_1})^*(\nu^{a_2})^* \ldots (\nu^{a_s})^*$$

со сложностью не выше n. Тогда его можно представить регулярным выражением правильного линейного вида с правильной линейной сложностью не выше $\frac{n^2}{4} + n$.

Теорема 2. *Пусть регулярный язык задаётся регулярным выражением следующего линейного вида*

$$A = \beta_1^* \beta_2^* \ldots \beta_s^*$$

со сложностью не выше n. Тогда его можно представить регулярным выражением правильного линейного вида с правильной линейной сложностью не выше $\frac{n^2}{2} + n$.

3. Вспомогательные утверждения

Лемма 1. *Пусть $x_1, \ldots, x_k$ — натуральные числа и $r :=$ НОД$(x_1, \ldots, x_k)$, $H := \{a_1 x_1 + \cdots + a_k x_k \mid a_1, \ldots, a_k \in \mathbb{N}_0\}$. Тогда существует $n_0 \in \mathbb{N}$ такое, что $Z_{n_0,r} \subseteq H$.*

Доказательство. Доказательство леммы приведено в главе 2 работы [1]. □

Лемма 2. *Пусть A — конечный алфавит и P — выражение линейного вида сложности не выше n. Тогда его можно представить выражением правильного линейного вида сложности не выше n^2.*

Доказательство. Данная лемма взята из главы 5 работы [1], там она приводится как лемма 3. □

Лемма 3. *Пусть A — конечный алфавит, $\alpha, \beta \in A^* \setminus \{\Lambda\}$ и $\alpha^k = \beta^m$ для некоторых $k, m \in \mathbb{N}$. Тогда существует $\nu \in A^* \setminus \{\Lambda\}$ такое, что*

$$|\nu| = \textit{НОД}(|\alpha|, |\beta|), \alpha = \nu^{\frac{|\alpha|}{|\nu|}}, \beta = \nu^{\frac{|\beta|}{|\nu|}}.$$

Доказательство. Доказательство леммы приведено в главе 2 работы [1]. $\square$

Лемма 4. *Пусть $a, b \in \mathbb{N}$, $a + b \leq n$. При таких ограничениях максимальное значение произведения ab не превосходит $\frac{n^2}{4}$.*

Доказательство. Рассмотрим два случая — когда сумма чётна и нечётна.

1) Пусть $a + b$ — чётно. Тогда для доказательства этой леммы воспользуемся неравенством между средним арифметическим и геометрическим:
$$\frac{a+b}{2} \geq \sqrt{ab},$$
где a и b — неотрицательные числа. Из этого неравенства следует, что
$$ab \leq \left(\frac{a+b}{2}\right)^2 = \frac{n^2}{4}.$$

2) Пусть $a + b$ — нечётно. Если a и b отличаются хотя бы на 2 и, без ограничения общности, $a \leq b$, то мы можем заменить a на $a+1$, а b — на $b-1$. От этого значение произведения может только вырасти. Таким образом, максимальное значение произведения достигается, когда $a = k, b = k+1$. И окончательно получаем
$$k(k+1) \leq \frac{(2k+1)^2}{4} = \frac{n^2}{4},$$
так как
$$4k^2 + 4k \leq 4k^2 + 4k + 1.$$

$\square$

Лемма 5. *Пусть $A = (0^{a_1})^*(0^{a_2})^* \dots (0^{a_s})^*$, $a_1 + a_2 + \dots + a_s \leq l$ и a_1, a_2 — взаимно просты. Тогда $T_{\textit{предп}}(X_0(A)) \leq a_1 a_2$.*

Доказательство. Рассмотрим следующие числа:

$$a_1,\ 2a_1,\ 3a_1\ \dots, a_2 a_1.$$

Они все входят в $X_0(A)$. Рассмотрим остатки этих чисел по модулю a_2. Из взаимной простоты a_1 и a_2 следует, что все эти остатки различны. А значит в этом списке встретятся все остатки. Самое большое число здесь — a_2a_1. Поэтому у любого другого числа $t \geq a_2a_1$ будет остаток, который мы уже получили ранее.

А раз у нас уже есть в списке число с таким остатком, то мы можем прибавить к нему a_2 несколько раз и получить число t. То есть, начиная с момента a_2a_1 период для X будет равен единице. А это значит, что $T_{\text{предп}}(X_0(A)) \leq a_1a_2$. □

Лемма 6. *Пусть* $A = (\nu^{a_1})^*(\nu^{a_2})^* \dots (\nu^{a_s})^*, s \geq 2$ *и выполнено неравенство* $(a_1+a_2+\dots+a_s)|\nu| \leq l$. *Тогда, для множества* $X_\nu(A) = \{l \mid \nu^l \in A\}$ *верно, что*

$$T_{\text{предп}}(X_\nu(A)) \leq \frac{l^2}{4|\nu|}.$$

Доказательство. Сначала докажем утверждение для случая, когда система чисел $a_1, a_2, \dots a_s$ взаимно проста в совокупности. Проведем доказательство методом обобщенной математической индукции по параметру s.

База индукции $(s = 2)$.

$$A = (\nu^{a_1})^*(\nu^{a_2})^*.$$

Рассмотрим вспомогательный случай $B = (0^{a_1})^*(0^{a_2})^*$.
Мы имеем $a_1 + a_2 \leq \frac{l}{|\nu|}$. Тогда по лемме 5 получаем

$$T_{\text{предп}}(X_0(B)) \leq a_1a_2.$$

И по лемме 4 это число не превосходит $\frac{l^2}{4|\nu|^2}$. Поэтому

$$T_{\text{предп}}(X_\nu(A)) \leq T_{\text{предп}}(X_0(B))|\nu| \leq \frac{l^2}{4|\nu|}.$$

Переход индукции. $(2, 3, \dots, s-1 \Rightarrow s)$
Докажем для s. Мы имеем

$$A = (\nu^{a_1})^*(\nu^{a_2})^* \dots (\nu^{a_s})^*.$$

Без ограничения общности упорядочим $a_1, a_2, \dots, a_s$ следующим образом

$$a_1 \leq a_2 \leq \dots \leq a_{s-1} \leq a_s.$$

Рассмотрим вспомогательное множество $\hat{A} = (0^{a_1})^*(0^{a_2})^* \dots (0^{a_s})^*$.
Тогда для множества

$$B = (0^{a_1})^*(0^{a_2})^* \dots (0^{a_{s-1}})^*$$

предпериод $T_п$ множества $X_0(B)$ не превосходит $\frac{(a_1+a_2+\cdots+a_{s-1})^2}{4}$ по предположению индукции. Обозначим через d_1 наибольший общий делитель системы чисел $a_1, a_2, \ldots, a_{s-1}$. По лемме 5 мы знаем, что начиная с момента $T_п$ множество $X_0(B)$ имеет период d_1. Рассмотрим вспомогательное множество $C = (0^{d_1})^*(0^{a_s})^*$. Мы знаем, что d_1 и a_s взаимно просты, поэтому по предположению базы предпериод множества $X_0(C)$ не превосходит $\frac{(d_1+a_s)^2}{4}$. Значит предпериод множества $X_0(\hat{A})$ не превосходит суммы предпериодов $X_0(B)$ и $X_0(C)$:

$$\frac{(a_1+a_2+\cdots+a_{s-1})^2}{4} + \frac{(d_1+a_s)^2}{4}.$$

Докажем, что это выражение не превосходит $\frac{l^2}{4|\nu|}$.

Для доказательства этого неравенства достаточно доказать, что

$$\frac{(a_1+a_2+\cdots+a_{s-1})^2}{4} + \frac{(d_1+a_s)^2}{4} \leq \frac{(a_1+a_2+\cdots+a_{s-1}+a_s)^2}{4},$$

что равносильно следующему неравенству

$$d_1^2 + 2d_1a_s \leq 2a_s(a_1+\cdots+a_{s-1}).$$

Что, очевидно, верно, так как $d_1 \leq a_1 \leq a_2 \leq \cdots \leq a_s$ и $s-1 \leq 2$. Возвращаясь к исходному множеству A, получаем

$$T_{\text{предп}}(X_\nu(A)) \leq T_{\text{предп}}(X_0(\hat{A}))|\nu| \leq \frac{l^2}{4|\nu|}.$$

Переход индукции доказан.

Теперь рассмотрим случай, когда НОД$(a_1, a_2, \ldots, a_s) = d > 1$.

Обозначим $\hat{\nu} = \nu^d$ и представим наше множество A в виде:

$$A = (\hat{\nu}^{\frac{a_1}{d}})^*(\hat{\nu}^{\frac{a_2}{d}})^* \ldots (\hat{\nu}^{\frac{a_s}{d}})^*.$$

Для такого представления множества мы уже выше доказали, что

$$T_{\text{предп}}(X_{\hat{\nu}}(A)) \leq \frac{l^2}{4|\hat{\nu}|} = \frac{l^2}{4d|\nu|}.$$

Поэтому

$$T_{\text{предп}}(X_\nu(A)) \leq T_{\text{предп}}(X_{\hat{\nu}}(A))d \leq \frac{l^2}{4d|\nu|}d = \frac{l^2}{4|\nu|}.$$

Лемма доказана. $\square$

Лемма 7. *Пусть $A = \beta_1^*\beta_2^* \ldots \beta_s^*$. Если соседние пары β_i в этом выражении соизмеримы, то тогда все β_i соизмеримы в совокупности.*

Доказательство. Данное утверждение легко доказывается индукцией по s. При $s = 2$ утверждение очевидно. Докажем переход индукции. Пусть мы уже доказали утверждение для множества $\beta_1^*\beta_2^*\dots\beta_{s-1}^*$. Тогда по индукционному предположению

$$\beta_1^*\beta_2^*\dots\beta_s^* = (\nu^{a_1})^*(\nu^{a_2})^*\dots(\nu^{a_{s-1}})^*\beta_s^*.$$

Осталось заметить, что из соизмеримости β_{s-1}, β_s и леммы 3 следует соизмеримость ν, β_s. Значит у слов $\beta_1, \beta_2, \dots, \beta_s$ есть общее измельчение.

Лемма доказана. □

4. Доказательство основных утверждений

Теорема 1. *Пусть регулярный язык задаётся регулярным выражением следующего линейного вида*

$$A = (\nu^{a_1})^*(\nu^{a_2})^*\dots(\nu^{a_s})^*$$

со сложностью не выше n. Тогда его можно представить регулярным выражением правильного линейного вида с правильной линейной сложностью не выше $\frac{n^2}{4} + n$.

Доказательство. При $s = 1$ утверждение очевидно, поэтому далее будем считать, что $s \geq 2$.

Для доказательства этой теоремы необходимо разобраться в доказательстве леммы 3 главы 5 работы [1]. Там показано, что наше представление задаётся в виде периодического семейства степеней ν, которое может быть получено конечным объединением элементов из предпериода и периодической части с длиной периода НОД$(a_1, a_2, \dots, a_s)$.

Каждое множество из этого объединения уже имеет правильный линейный вид. Самое большое по длине из этих множеств задаётся в следующем виде:

$$\nu^{T_{\text{предп}}(X_\nu(A))}(\nu^{\text{НОД}(a_1,a_2,\dots,a_s)})^*. \qquad (*)$$

Оценим его длину. По лемме 6

$$T_{\text{предп}}(X_\nu(A)) \leq \frac{n^2}{4|\nu|}.$$

И НОД$(a_1, a_2, \dots, a_s) \leq a_1 \leq a_1 + \dots + a_s \leq \frac{n}{|\nu|}$.

Поэтому длина выражения (*) не превосходит

$$|\nu|(T_{\text{предп}}(X_\nu(A)) + \text{НОД}(a_1, a_2, \dots, a_s)) \leq |\nu|\left(\frac{n^2}{4|\nu|} + \frac{n}{|\nu|}\right) = \frac{n^2}{4} + n.$$

Теорема доказана. □

Теорема 2. *Пусть регулярный язык задаётся регулярным выражением следующего линейного вида*

$$A = \beta_1^* \beta_2^* \dots \beta_s^*$$

со сложностью не выше n. Тогда его можно представить регулярным выражением правильного линейного вида с правильной линейной сложностью не выше $\frac{n^2}{2} + n$.

Доказательство. Для доказательства данного утверждения нужно глубже погрузиться в доказательство леммы 3 главы 4 работы [1]. В ней рассматриваются процедура "расщепления", позволяющая преобразовывать соседнюю пару итераций $\beta_1^*\beta_2^*$ в выражение правильного линейного вида.

На идейном уровне эту процедуру можно описать так: из итерации β_2^* вытягивается несколько раз слово β_2 так, чтобы результат нельзя было целиком перебросить через итерацию β_1. И показывается, что длина вытянутого слова не превосходит $|\beta_1||\beta_2|$.

С помощью этого факта можно доказать утверждение нашей теоремы. Для этого разобьём выражение $\beta_1^*\beta_2^* \dots \beta_s^*$ в последовательность из блоков $B_1, B_2, \dots, B_k$, где слова в каждом из блоков B_i и B_{i+1} попарно соизмеримы, а значит по лемме 7 все блоки B_i представимы в виде

$$B_i = (\nu_i^{a_1(i)})^*(\nu_i^{a_2(i)})^* \dots (\nu_i^{a_s(i)})^*.$$

При этом ν_i и ν_{i+1} несоизмеримы в силу разбиения на блоки. Обозначим длину блока B_i через l_i. Из теоремы 1 знаем, что каждый блок B_i можно заменить на представление правильного линейного вида с длиной не больше чем $\frac{l_i^2}{4} + l_i$. При этом очень важно отметить, что в новом линейном виде все подъитерационные слова являются степенями слова ν_i.

Дальше в ход идёт процедура "расщепления" для пары несоизмеримых звёзд, позволяющая "развести" соседние блоки B_i. Для этого, как было показано в доказательстве теоремы 1 потребуется увеличить длину выражения на не более чем $l_1 l_2 + l_2 l_3 + \dots + l_{s-1} l_s$ символов.

Совмещая вместе идеи о "расщеплении" и приведении блоков в правильный линейный вид окончательно получаем оценку на длину нового линейного вида:

$$\frac{l_1^2}{4} + l_1 + \dots + \frac{l_s^2}{4} + l_s + (l_1 l_2 + l_2 l_3 + \dots + l_{s-1} l_s) =$$

$$= \frac{l_1^2 + l_2^2 + \dots + l_s^2 + 4l_1 l_2 + 4l_2 l_3 + \dots + 4l_{s-1} l_s}{4} + (l_1 + l_2 + \dots + l_s) \le$$

$$\leq \frac{(l_1 + l_2 + \cdots + l_s)^2}{4} + \frac{2l_1l_2 + 2l_2l_3 + \cdots + 2l_{s-1}l_s}{4} + (l_1 + l_2 + \cdots + l_s) \leq$$

$$\leq \frac{(l_1 + l_2 + \cdots + l_s)^2}{4} + \frac{(l_1 + l_2 + \cdots + l_s)^2}{4} + (l_1 + l_2 + \cdots + l_s) =$$

$$= \frac{(l_1 + l_2 + \cdots + l_s)^2}{2} + (l_1 + l_2 + \cdots + l_s).$$

Правильность полученного линейного вида следует из выше отмеченного факта о том, что подъитерационные слова в заменённых блоках по прежнему начинаются с ν_i и не портят условие правильности при использовании процедуры "расщепления".

Теорема доказана. □

5. Заключение

В данной статье было проведено исследование о представлении регулярных языков с полиномиальной функцией роста с использованием регулярных выражений линейного вида. Основной целью публикации было разработать новую оценку сложности перехода от линейного вида регулярных языков к правильному линейному виду.

В ходе исследования были рассмотрены существующие оценки сложности перехода и проведен анализ их эффективности. Было обнаружено, что существующие оценки не всегда достаточно точно оценивают сложность перехода. Это стало основной мотивацией для разработки новой оценки.

На основе проведенного исследования была сформулирована и доказана основная теорема, утверждающая, что регулярный язык, задаваемый регулярным выражением линейного вида с линейной сложностью не превышающей n, может быть задан регулярным выражением правильного линейного вида с линейной сложностью не превышающей $\frac{n^2}{2} + n$. Это означает, что предложенная новая оценка более точно оценивает сложность перехода и может быть применена для оптимизации представления регулярных языков.

Список литературы

[1] Дергач П.С., "Алфавитное кодирование регулярных языков с полиномиальной функцией роста", *Кандидатская диссертация*, 2016, 1–213.

[2] Абдурахмонов А.Н., "О правильной линейной форме", *Выпускная квалификационная работа*, 2019, 1–23.

[3] В. Б. Кудрявцев, С. В. Алешин, А. С. Подколзин., “Введение в теорию автоматов”, *М.: Наука*, 1985.

[4] Яблонский С.В., “Введение в дискретную математику”, *М.: Наука*, 1986.

[5] Виноградов И.М., “Введение в теорию чисел”, *М.: Наука*, 1972.

On the complexity of converting to a correct linear form

Dergach P.S., Saltsova D.A.

This paper is devoted to the study of the regular linear form for regular languages with polynomial growth function and improving the corresponding estimates on the complexity of the transition from the linear form to the regular linear form.

We managed to lower the previously known estimate n^2 from [1] to an estimate $\frac{n^2}{2} + n$. Also obtained an upper estimate $\frac{n^2}{4}$ for languages of the form $\beta_1^*\beta_2^* \ldots \beta_s^*$, in which the words β_i are commensurable.

Keywords: regular language, growth function, regular linear form, complexity.

References

[1] Dergach P.S., “Alphabetic encoding of regular languages with polynomial growth function”, *PhD thesis*, 2016, 1–213.

[2] Abdurakhmonov A.N., “About the correct linear form”, *bachelor thesis*, 2019, 1–23.

[3] Kudryavtsev V.B., Aleshin S.V, Podkolzin A.S., “Introduction to automata theory”, 1985.

[4] Yablonskiy S.V., “Introduction to discrete math”, 1986.

[5] Vinogradov I. M., “Basics of number theory”, 1972.