

МАТЕМАТИЧЕСКАЯ ЛОГИКА, АЛГЕБРА, ТЕОРИЯ ЧИСЕЛ И ДИСКРЕТНАЯ МАТЕМАТИКА

УДК 510.67, 519.766

ОПРЕДЕЛИМОСТЬ МНОЖЕСТВ В ТЕОРИЯХ ЯЗЫКОВ С ОПЕРАЦИЯМИ ОБЪЕДИНЕНИЯ И ИТЕРАЦИИ

Карлов Б.Н.

Тверской государственный университет, г. Тверь

Поступила в редакцию 30.05.2025, после переработки 24.06.2025.

В работе исследованы теории языков в различных алфавитах с операциями объединения и итерации. Доказано, что в случае однобуквенного алфавита в такой теории определима операция возведения языка, содержащего единственное слово, в любую фиксированную степень. Как следствие, установлена определимость всех конечных языков и всех языков с конечными дополнениями. Доказано, что в случае многобуквенного алфавита неопределимы некоторые конечные языки. Установлено, что, независимо от алфавита, определимы языки, не являющиеся контекстно-свободными.

Ключевые слова: формальный язык, теория, объединение, итерация, определимость.

Вестник ТвГУ. Серия: Прикладная математика. 2025. № 2. С. 22–31.
<https://doi.org/10.26456/vtpmk734>

Введение

Одной из задач теории моделей является исследование определимости, то есть возможности выразить одни отношения и функции через другие с помощью формул некоторой логики. Важным частным случаем является возможность выделения фиксированных элементов основного множества. Математическая логика вообще и определимость в частности тесно связаны с теорией автоматов и формальных языков. Классическим результатом является теорема Бюхи–Элгота–Трахтенброта [1, 2, 4], утверждающая, что класс регулярных языков совпадает с классом языков, определимых в монадической логике второго порядка. Однако можно рассматривать определимость языков не в общей логике, а в некоторых теориях, в частности в теориях различных семейств языков.

Если для какого-либо класса языков зафиксировать набор операций, относительно которых этот класс замкнут, то получится алгебра, в которой основным множеством является соответствующий класс языков. Можно рассматривать алгебры всех языков, конечных языков, регулярных или контекстно-свободных языков и т. п. и можно изучать возможность определений тех или иных операций

над языками в получающихся теориях. В работе [3] были исследованы некоторые теории класса регулярных языков в однобуквенном алфавите, в частности теория языков с операциями объединения и итерации. Было доказано, что эта теория алгоритмически эквивалентна элементарной арифметике и, следовательно, неразрешима. Для доказательства была установлена определимость отношений $\text{Mult}_k(x, y)$ для $k \in \{2, 3, 5, 7\}$, означающих, что $x = \{a^n\}$ для некоторого n , $y = \{a^{n^k}\}$. Это позволило закодировать конфигурации счётчиковых машин в показательном виде и интерпретировать сложение и умножение натуральных чисел. Используемый для определения Mult_k метод применим и для класса всех языков. Однако оставался открытым вопрос о возможности определения этих отношений для произвольных k .

В настоящей статье мы продолжаем изучать теории языков с операциями объединения и конкатенации в произвольном алфавите. В разделе 2 мы доказываем, что в случае однобуквенного алфавита отношения Mult_k определимы для любого $k \geq 0$. В качестве следствия мы получаем определимость всех конечных языков и всех языков с конечными дополнениями. В разделе 3 мы рассматриваем многобуквенные алфавиты и доказываем, что в этом случае неопределим даже язык $\{ab\}$. Однако, несмотря на такие ограниченные возможности, даже при наличии только одного символа определимы языки, не являющиеся контекстно-свободными.

1. Предварительные сведения

Мы будем рассматривать множество всех языков в некотором фиксированном алфавите. Теоретико-множественные операции определяются для языков как для произвольных множеств. *Итерацией* языка L называется множество слов, представимых в виде конкатенации нескольких слов из L :

$$L^* = \{w_1 w_2 \dots w_n : n \geq 0, w_i \in L\}.$$

В частности, пустое слово ε принадлежит языку L^* для любого L . По определению $\emptyset^* = \{\varepsilon\}$. Для произвольного слова w и натурального числа n через w^n обозначается n -я степень слова w : $w^n = \underbrace{ww \dots w}_{n \text{ раз}}$. В частности, $w^0 = \varepsilon$.

Теорией называется множество замкнутых формул первого порядка, замкнутое относительно выводимости. Формулы φ и ψ называются *эквивалентными* в теории T , если $\varphi \leftrightarrow \psi \in T$. Эквивалентность формул в теории T обозначается $\varphi \equiv_T \psi$. Отношение $R(x_1, \dots, x_n)$ *определимо* в теории T , если существует формула $\varphi(x_1, \dots, x_n)$, не содержащая R , такая, что

$$(\forall x_1) \dots (\forall x_n)(R(x_1, \dots, x_n) \leftrightarrow \varphi(x_1, \dots, x_n)) \in T.$$

Функция $f(x_1, \dots, x_n)$ *определима* в теории T , если существует такая формула $\varphi(y, x_1, \dots, x_n)$, не содержащая f , что

$$(\forall y)(\forall x_1) \dots (\forall x_n)(y = f(x_1, \dots, x_n) \leftrightarrow \varphi(y, x_1, \dots, x_n)) \in T.$$

В частности, константа c *определима*, если для некоторой формулы $\varphi(x)$, не содержащей c , имеет место

$$(\forall x)(x = c \leftrightarrow \varphi(x)) \in T.$$

Пусть $\Sigma_m = \{a_1, a_2, \dots, a_m\}$ — произвольный алфавит ($m \geq 1$), и пусть $\mathfrak{L}_m$ — некоторый класс языков в алфавите Σ_m такой, что:

- $\mathfrak{L}_m$ образует булеву алгебру с наименьшим элементом $\emptyset$ и наибольшим элементом Σ_m^* (множество всех слов в алфавите Σ_m);
- $\mathfrak{L}_m$ замкнут относительно итерации;
- $\mathfrak{L}_m$ содержит все конечные языки.

Обозначим через $\mathfrak{A}_m$ алгебру с основным множеством $\mathfrak{L}_m$ сигнатуры

$$\Omega_m = (\cup^{(2)}, *^{(1)}, a_1^{(0)}, \dots, a_m^{(0)}).$$

Операции $\cup$ и $*$ интерпретируются как объединение и итерация соответственно, а константы a_i обозначают языки $\{a_i\}$. Через T_m обозначим теорию алгебры $\mathfrak{A}_m$, то есть множество всех формул первого порядка, истинных в алгебре $\mathfrak{A}_m$. Поскольку $\mathfrak{L}_m$ образует булеву алгебру, то для любого m в теории T_m определимы операции пересечения $\cap$ и разности $\setminus$, константы $\emptyset$ и Σ_m^* , обозначающие пустой язык и множество всех слов соответственно, а также отношение включения $\subseteq$. Кроме того, для любого натурального $i \geq 0$ определимо отношение $\text{Card}_i(x)$, означающее, что множество x имеет мощность i . Мы будем также использовать константу ε , обозначающую язык $\{\varepsilon\}$. Она определима как $\emptyset^*$.

2. Определимость в случае однобуквенного алфавита

В этом разделе мы будем рассматривать теорию T_1 . Для краткости мы будем опускать индекс у символа a_1 и считать, что $\Sigma_1 = \{a\}$. Обозначим через $\text{Mult}_k(x, y)$ следующее отношение: $x = \{a^n\}$ для некоторого $n \geq 0$, $y = \{a^{kn}\}$. Для доказательства определмости этих отношений в теории T_1 мы будем использовать следующее свойство итерации.

Лемма 1. Пусть $L_{n,k} = \{\varepsilon\} \cup \{a^{in} : i \geq k\} = \{\varepsilon, a^{kn}, a^{(k+1)n}, a^{(k+2)n}, \dots\}$ для некоторых $n > 0$, $k \geq 0$. Тогда:

- $L_{n,k}$ является итерацией языка $L = \{a^{kn}, a^{(k+1)n}, \dots, a^{(2k-1)n}\}$, содержащего ровно k слов;
- $L_{n,k}$ не является итерацией никакого другого языка, содержащего не более k слов.

Доказательство. Равенство $L^* = L_{n,k}$ проверяется непосредственно. Язык L^* содержит все слова $a^{kn}, \dots, a^{(2k-1)n}$, так как они принадлежат L , а уже слово a^{2kn} может быть представлено в виде конкатенации $a^{kn}a^{kn}$. При этом L^* не содержит других слов, поскольку длины всех слов из L кратны n и в L нет слов, длина которых меньше чем kn . Наоборот, если $L_1^* = L_{n,k}$ для некоторого L_1 , то L_1 должен содержать все слова из L , так как ни одно из них не может быть получено как конкатенация более коротких слов. $\square$

Теперь мы можем доказать наш основной результат.

Теорема 1. *Отношения Mult_k определены в теории T_1 для любого $k \geq 0$.*

Доказательство. Индукция по k .

Базис индукции. Для $k = 0$ или $k = 1$ отношения определяются следующим образом:

$$\begin{aligned}\text{Mult}_0(x, y) &\equiv_{T_1} y = \varepsilon, \\ \text{Mult}_1(x, y) &\equiv_{T_1} x = y.\end{aligned}$$

Определимость отношения $\text{Mult}_2(x, y)$ была доказана в [3].

Индукционный шаг. Рассмотрим два случая в зависимости от чётности k . Предположим сначала, что k чётно, то есть $k = 2p$ для некоторого p . По индукционному предположению отношение Mult_p определимо. Тогда

$$\text{Mult}_k(x, y) \equiv_{T_1} (\exists z)(\text{Mult}_p(x, z) \wedge \text{Mult}_2(z, y)).$$

Действительно, если $x = \{a^n\}$, $y = \{a^{2pn}\}$, то можно положить $z = \{a^{pn}\}$. Наоборот, по индукционному предположению из истинности $\text{Mult}_p(x, z)$ следует, что $x = \{a^n\}$, $z = \{a^{pn}\}$ для некоторого n , а значит, из истинности $\text{Mult}_2(z, y)$ следует, что $y = \{a^{2pn}\} = \{a^{kn}\}$.

Пусть теперь k нечётно, то есть $k = 2p + 1$ для некоторого p . Для произвольных чисел s и t таких, что $s \leq t$, обозначим через $\text{Mult}_{[s,t]}(x, y)$ следующее отношение: $x = \{a^n\}$ для некоторого n , $y = \{a^{in} : s \leq i \leq t\}$. Тогда

$$\text{Mult}_{[s,t]}(x, y) \equiv_{T_1} (\forall z) \left((\text{Card}_1(z) \wedge z \subseteq y) \leftrightarrow \bigvee_{i=s}^t \text{Mult}_i(x, z) \right).$$

Если $t \leq p$, то все отношения Mult_i в дизъюнкции определены по индукционному предположению, а значит, определимо и отношение $\text{Mult}_{[s,t]}$. Отношение Mult_k определяется следующим образом:

$$\begin{aligned}\text{Mult}_k(x, y) &\equiv_{T_1} ((x = \varepsilon \wedge y = \varepsilon) \vee \\ &\vee (\text{Card}_1(x) \wedge x \neq \varepsilon \wedge (\exists z)(\exists t)(\exists u)(\exists v)(\text{Mult}_{[1,2p]}(x, z) \wedge \text{Card}_{2p+1}(u) \wedge \\ &\wedge x^* \setminus z = u^* \wedge \text{Mult}_{[1,p]}(x, t) \wedge \text{Card}_{p+1}(v) \wedge x^* \setminus t = v^* \wedge y = u \cap v))).\end{aligned}$$

Докажем правильность этого определения. Предположим, что $x = \{a^n\}$ для некоторого n , $y = \{a^{kn}\}$, и докажем, что в этом случае формула в правой части истинна. Если $n = 0$, то истинен первый член дизъюнкции. Пусть $n > 0$. Формула $\text{Card}_1(x)$ истинна. Выберем для переменных u, v, z, t следующие значения:

$$\begin{aligned}u &= \{a^{(2p+1)n}, a^{(2p+2)n}, \dots, a^{(4p+1)n}\}, & z &= \{a^n, a^{2n}, \dots, a^{2pn}\}, \\ v &= \{a^{(p+1)n}, \dots, a^{(p+2)n}, \dots, a^{(2p+1)n}\}, & t &= \{a^n, a^{2n}, \dots, a^{pn}\}.\end{aligned}$$

Языки u и v содержат $2p + 1$ и $p + 1$ слов соответственно, поэтому формулы $\text{Card}_{2p+1}(u)$ и $\text{Card}_{p+1}(v)$ истинны, а по индукционному предположению истинны также формулы $\text{Mult}_{[1,2p]}(x, z)$ и $\text{Mult}_{[1,p]}(x, t)$. Итерация языка x равна $x^* = \{\varepsilon, a^n, a^{2n}, a^{3n}, \dots\}$, поэтому

$$x^* \setminus z = \{\varepsilon, a^{(2p+1)n}, a^{(2p+2)n}, \dots\}, \quad (1)$$

$$x^* \setminus t = \{\varepsilon, a^{(p+1)n}, a^{(p+2)n}, \dots\}. \quad (2)$$

По лемме 1 справедливы равенства $x^* \setminus z = u^*$ и $x^* \setminus t = v^*$. Поскольку кроме того $u \cap v = \{a^{(2p+1)n}\} = y$, то вся формула в правой части истинна.

Наоборот, предположим, что формула истинна. Если истинен первый член дизъюнкции, то $x = y = \{\varepsilon\} = \{a^0\}$. Если истинен второй член, то $x = \{a^n\}$ для некоторого n в силу $\text{Card}_1(x)$. Следовательно, по индукционному предположению обязательно

$$z = \{a^n, a^{2n}, \dots, a^{2pn}\}, \quad t = \{a^n, a^{2n}, \dots, a^{pn}\},$$

поэтому языки $x^* \setminus z$ и $x^* \setminus t$ определяются равенствами (1) и (2). Поскольку $\text{Card}_{2p+1}(u)$ и $\text{Card}_{p+1}(v)$ истинны, то по лемме 1 равенства $x^* \setminus z = u^*$, $x^* \setminus t = v^*$ возможны только при

$$\begin{aligned} u &= \{a^{(2p+1)n}, a^{(2p+2)n}, \dots, a^{(4p+1)n}\}, \\ v &= \{a^{(p+1)n}, \dots, a^{(p+2)n}, \dots, a^{(2p+1)n}\}. \end{aligned}$$

Но тогда $y = u \cap v = \{a^{(2p+1)n}\} = \{a^{kn}\}$. □

Отметим несколько простых следствий.

Следствие 1. Для любого n в теории T_1 определим язык $\{a^n\}$.

Доказательство. Язык $\{a^n\}$ выделяется формулой $\text{Mult}_n(a, x)$. □

Следствие 2. Для любого $n \geq 1$ в теории T_1 определим язык $L = \{a^{nk} : k \geq 0\}$.

Доказательство. Язык L является итерацией языка $\{a^n\}$ из предыдущего следствия. □

Заметим, что в теории T_1 определимы и некоторые другие арифметические прогрессии. Например, определимо множество всех слов нечётной длины как разность $a^* \setminus (a^2)^*$.

Следствие 3. В теории T_1 определимы все конечные языки и все языки с конечными дополнениями.

Доказательство. Конечный язык $L = \{a^{n_1}, a^{n_2}, \dots, a^{n_k}\}$ представляется в виде объединения $\{a^{n_1}\} \cup \dots \cup \{a^{n_k}\}$, а его дополнение — в виде разности $a^* \setminus L$. □

В частности, эти результаты справедливы для алгебр, основным множеством которых являются все языки, все регулярные, контекстно-зависимые или рекурсивные языки.

3. Определимость в случае многобуквенного алфавита

Если алфавит содержит более одного символа, то следствие 3 уже может не выполняться. Даже некоторые одноэлементные языки не определимы в таких теориях, если носитель соответствующей алгебры удовлетворяет некоторым дополнительным ограничениям.

Обозначим через w^R обращение слова w : если $w = a_1 \dots a_n$, то $w^R = a_n \dots a_1$. Операция обращения поэлементно переносится на языки: $L^R = \{w^R : w \in L\}$.

Теорема 2. Если $i \geq 2$, $a, b \in \Sigma_i$ — два различных символа и класс языков $\mathfrak{L}_i$ замкнут относительно обращения, то в теории T_i не определим язык $\{ab\}$.

Доказательство. Докажем, что операция обращения является автоморфизмом алгебры языков $\mathfrak{A}_i$. Равенство $(L_1 \cup L_2)^R = L_1^R \cup L_2^R$ получается из следующих эквивалентностей:

$$\begin{aligned} w \in (L_1 \cup L_2)^R &\iff w^R \in L_1 \cup L_2 \iff w^R \in L_1 \text{ или } w^R \in L_2 \iff \\ &\iff w \in L_1^R \text{ или } w \in L_2^R \iff w \in L_1^R \cup L_2^R. \end{aligned}$$

Докажем равенство $(L^*)^R = (L^R)^*$. Если $w \in (L^*)^R$, то $w^R \in L^*$, а значит, $w^R = w_1 w_2 \dots w_n$ для некоторого n и некоторых слов $w_i \in L$. Поэтому $w = w_n^R \dots w_2^R w_1^R$, причём $w_i^R \in L^R$. Следовательно, $w \in (L^R)^*$. Доказательство в другую сторону симметрично.

Поскольку $\{ab\}^R = \{ba\}$, то мы нашли автоморфизм, не сохраняющий константу $\{ab\}$. Отсюда следует, что эта константа не определима в теории T_i (см. [5]). $\square$

Следствие 4. Если $i \geq 2$ и класс языков $\mathfrak{L}_i$ замкнут относительно обращения, то в теории T_i не определима конкатенация.

Доказательство. Если бы формула $\varphi(x, y, z)$ выражала равенство $xy = z$, то формула $\varphi(a, b, z)$ определяла бы константу $\{ab\}$. $\square$

Поскольку замкнутость относительно обращения имеет место для классов всех языков, регулярных, контекстно-зависимых и рекурсивных языков, то в соответствующих теориях не определима конкатенация.

Итак, мы показали, что в теориях T_i не определимы даже многие конечные языки. Однако в них определимы достаточно сложные языки. Напомним, что слово w называется *примитивным*, если w непусто и не существует слова v и числа $n \geq 2$ таких, что $v^n = w$ (см. [6]). Неизвестно, является ли множество всех примитивных слов контекстно-свободным в случае алфавита, содержащего хотя бы два символа. Однако оно определимо в теориях T_i .

Теорема 3. Если $\mathfrak{L}_i$ — множество всех языков в алфавите Σ_i , то для любого i в теории T_i определимо множество всех примитивных слов.

Доказательство. Следующая формула Primitive_1 выделяет языки мощности 1, содержащие примитивное слово:

$$\text{Primitive}_1(x) \equiv_{T_i} x \neq \varepsilon \wedge \text{Card}_1(x) \wedge (\forall y)((\text{Card}_1(y) \wedge x \subseteq y^*) \rightarrow x = y).$$

Пусть $x = \{w\}$, где w примитивно. Рассмотрим произвольный одноэлементный язык $y = \{v\}$ такой, что $x \subseteq y^*$, и докажем, что заключение импликации истинно. Так как $x \subseteq y^*$ и $x \neq \{\varepsilon\}$, то $w = v^n$ для некоторого $n \geq 1$. Но случай $n > 1$ невозможен в силу примитивности w . Следовательно, $n = 1$, $w = v$ и $x = y$.

Наоборот, предположим, что x не имеет вида $\{w\}$ ни для какого примитивного слова w . Если x содержит больше одного слова или равен $\{\varepsilon\}$, то ложен первый или второй член конъюнкции. Если $x = \{v^n\}$ для $v \neq \varepsilon$, $n > 1$, то положим

$y = \{v\}$. Так как y^* содержит слово v^n , то посылка импликации истинна, но при этом заключение ложно, а значит, и в этом случае вся формула ложна.

Теперь множество всех примитивных слов определяется следующим образом:

$$\text{Primitive}(x) \equiv_{T_1} (\forall y)((\text{Card}_1(y) \wedge y \subseteq x) \leftrightarrow \text{Primitive}_1(y)). \quad \square$$

Немного изменив формулу, можно определить и язык, не являющийся контекстно-свободным.

Теорема 4. Если $\mathfrak{L}_i$ — множество всех языков в алфавите Σ_i , то для любого $i \geq 1$ и для любого $a \in \Sigma_i$ в теории T_i определим язык

$$L_{\text{prime}} = \{a^n : n \text{ — простое число}\}.$$

Доказательство. Пусть сначала $i = 1$. Следующая формула Prime_1 выделяет языки мощности 1, содержащие слово, длина которого является простым числом:

$$\begin{aligned} \text{Prime}_1(x) \equiv_{T_1} & \text{Card}_1(x) \wedge x \neq \varepsilon \wedge x \neq a \wedge \\ & \wedge (\forall y)((\text{Card}_1(y) \wedge y \neq a \wedge x \subseteq y^*) \rightarrow x = y). \end{aligned}$$

Аналогично формуле $\text{Primitive}_1(x)$ из доказательства предыдущей теоремы, формула $\text{Prime}_1(x)$ говорит, что x не является степенью никакого слова отличного от a . Тогда множество слов простой длины определяется следующим образом:

$$\text{Prime}(x) \equiv_{T_1} (\forall y)((\text{Card}_1(y) \wedge y \subseteq x) \leftrightarrow \text{Prime}_1(y)).$$

Если $i \geq 2$, то нужно построить релятивизацию формулы $\text{Prime}(x)$, заменив подформулы вида $(\exists z)\varphi(z)$ на $(\exists z)(z \subseteq a^* \wedge \varphi(z))$, а подформулы вида $(\forall z)\varphi(z)$ — на $(\forall z)(z \subseteq a^* \rightarrow \varphi(z))$. $\square$

Заключение

В работе была исследована возможность определения различных языков в теории языков с операциями объединения и итерации. Была доказана возможность определения любого языка вида $\{a^{nk} : n \geq 0\}$ для любой константы k , а также всех конечных языков и языков с конечными дополнениями в случае однобуквенного алфавита. Для многобуквенных алфавитов был приведён пример неопределимого языка из одного слова и пример определимого языка, не являющегося контекстно-свободным. Сформулируем некоторые вопросы для дальнейшего исследования.

- Определимы ли в теории T_1 все языки вида $\{a^{kn+l} : n \geq 0\}$, где k и l — фиксированные натуральные числа? Из утвердительного ответа на этот вопрос будет следовать, что в теории T_1 определимы все регулярные языки.
- Определимы ли в теориях T_i при $i \geq 2$ все языки вида $\{w \in \Sigma^* : |w| = n\}$, где n — фиксированная длина?

Список литературы

- [1] Трахтенброт Б.А. Конечные автоматы и логика одноместных предикатов // Сибирский математический журнал. 1962. Т. 3, № 1. С. 103–131.
- [2] Büchi J.R. Weak second order arithmetic and finite automata // Zeitschrift für mathematische Logik und Grundlagen der Mathematik. 1960. Vol. 6. Pp. 66–92.
- [3] Dudakov S., Karlov B. On decidability of theories of regular languages // Theory of Computing Systems. 2021. Vol. 65. Pp. 462–478. <http://doi.org/10.1007/s00224-020-09995-4>
- [4] Elgot C.C. Decision problems of finite automata design and related arithmetics // Transactions of the American Mathematical Society. 1961. Vol. 98. Pp. 21–52. <https://doi.org/10.1090/S0002-9947-1961-0139530-9>
- [5] Marker D. Model theory: an introduction. New York: Springer-Verlag, 2002. 345 p.
- [6] Shallit J. A Second Course in Formal Languages and Automata Theory. Cambridge: Cambridge University Press, 2008. 240 p.

Образец цитирования

Карлов Б.Н. Определимость множеств в теориях языков с операциями объединения и итерации // Вестник ТвГУ. Серия: Прикладная математика. 2025. № 2. С. 22–31. <https://doi.org/10.26456/vtpmk734>

Сведения об авторах

1. Карлов Борис Николаевич

доцент кафедры информатики факультета прикладной математики и кибернетики Тверского государственного университета.

Россия, 170100, г. Тверь, ул. Желябова, 33, ТвГУ. E-mail: bnkarlov@gmail.com

DEFINABILITY OF SETS IN THEORIES OF LANGUAGES WITH UNION AND KLEENE STAR

Karlov B.N.

Tver State University, Tver

Received 30.05.2025, revised 24.06.2025.

In this paper we study theories of languages in different alphabets with the operations of union and Kleene star. We prove that in the case of one-symbol alphabet such theory allows to define the power operation on singleton languages for an arbitrary fixed exponent. As corollary we establish definability of all finite and cofinite languages. It is proved that in the case of multi-symbol alphabets some finite languages are undefinable. It is established that independently on the alphabet some non-context-free languages are definable.

Keywords: formal language, theory, union, Kleene star, definability.

Citation

Karlov B.N., “Definability of sets in theories of languages with union and Kleene star”, *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2025, № 2, 22–31 (in Russian). <https://doi.org/10.26456/vtpmk734>

References

- [1] Trakhtenbrot B.A., “Finite automata and the logic of one-place predicates”, *Sibirskij matematicheskij zhurnal [Siberian Mathematical Journal]*, **3**:1 (1962), 103–131 (in Russian).
- [2] Büchi J.R., “Weak second order arithmetic and finite automata”, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **6** (1960), 66–92.
- [3] Dudakov S., Karlov B., “On decidability of theories of regular languages”, *Theory of Computing Systems*, **65** (2021), 462–478, <http://doi.org/10.1007/s00224-020-09995-4>.
- [4] Elgot C.C., “Decision problems of finite automata design and related arithmetics”, *Transactions of the American Mathematical Society*, **98** (1961), 21–52, <https://doi.org/10.1090/S0002-9947-1961-0139530-9>.
- [5] Marker D., *Model theory: an introduction*, Springer-Verlag, New York, 2002, 345 pp.
- [6] Shallit J., *A Second Course in Formal Languages and Automata Theory*, Cambridge University Press, Cambridge, 2008, 240 pp.

Author Info**1. Karlov Boris Nikolaevich**

Associate Professor at Computer Science department, Faculty of Applied Mathematics and Cybernetics, Tver State University.

Russia, 170100, Tver, 33 Zhelyabova str., TverSU. E-mail: bnkarlov@gmail.com