

УДК 519.713.2+512.563

ВЫЧИСЛИМЫЕ ИЗОМОРФИЗМЫ ОТНОСИТЕЛЬНО РЕГУЛЯРНЫХ БУЛЕВЫХ АЛГЕБР

И. Н. Шиманогов, М. Н. Вялый

Аннотация. Рассматривается класс булевых алгебр, образуемых пересечениями регулярных языков с заданным языком. В случае, когда такая алгебра изоморфна алгебре регулярных языков, доказывается существование изоморфизма, вычислимого с использованием оракулов для задачи регулярной реализуемости и бесконечной регулярной реализуемости. Данный результат дает существование вычислимого изоморфизма булевых алгебр регулярных языков над алфавитом из одной буквы и из двух букв. Также строится нижняя оценка сложности этой задачи.

DOI 10.33048/smzh.2026.67.519

Ключевые слова: задачи регулярной реализуемости, булевы алгебры, относительно регулярные алгебры, вычислительная сложность, регулярные языки.

1. Введение

Регулярные языки являются одним из ключевых объектов теории формальных языков. Важным классом алгоритмических задач является проверка существования в фиксированном языке слова, удовлетворяющего некоторому регулярному условию. Данные задачи называются задачами регулярной реализуемости и были введены М. Н. Вялым в [1]. Формальная постановка задачи следующая. Рассмотрим произвольный язык L , входом является регулярный язык R , заданный, например, некоторым автоматом. Необходимо проверить непустоту пересечения входа R и фиксированного языка L .

Задачи регулярной реализуемости для контекстно-свободных языков активно изучались под названием «задачи контекстно-свободной достижимости», см., например [2, 3]. В [4] изучаются приложения этой задачи к теории тонких сводимостей. Разрешимость задач регулярной реализуемости доказана для широкого круга графовых задач [5–7], а также для задач, связанных с целочисленным программированием [8].

В данной статье мы работаем со следующим алгебраическим обобщением задачи регулярной реализуемости. По любому языку L можно построить булеву алгебру, состоящую из всевозможных пересечений регулярных языков с L и операциями объединения, пересечения и дополнения до L . Данные булевы алгебры называются *относительно регулярными алгебрами*. Оказывается, что класс алгебр такого вида весьма широк: для любой счетной атомной булевой алгебры существует изоморфная ей относительно регулярная алгебра [9].

Исследование второго автора финансировалось в рамках госзадания FFNG-2024-0003.

При этом существует естественный гомоморфизм алгебры регулярных языков на относительно регулярную алгебру языка L : $\varphi(R) = R \cap L$. Таким образом, задача регулярной реализуемости является задачей проверки того, что данный элемент алгебры (регулярный язык) не лежит в ядре этого гомоморфизма.

Данная работа посвящена случаю, когда относительно регулярная алгебра изоморфна алгебре регулярных языков. Оказывается, что подобный случай охватывает относительно регулярные алгебры довольно широкого класса языков, в том числе и все те, которые порождаются контекстно-свободными языками. В работе показывается, что при условии разрешимости задач регулярной реализуемости и бесконечной регулярной реализуемости данный изоморфизм оказывается вычислимым.

Полученный результат применим к изоморфизму между булевыми алгебрами регулярных языков над алфавитом из одной и двух букв. Для этого случая дается нижняя оценка на вычислительную сложность.

2. Формальные языки

Напомним определения и сформулируем теоремы теории формальных языков, которые будут использоваться в данной работе. Далее будем называть *натуральными* числами целые неотрицательные числа и обозначать множество таких чисел символом $\mathbb{N}$. *Языком* будем называть множество слов в конечном алфавите. Так как свойства языков зависят только от размера алфавита, то, без ограничения общности и если не оговорено иное, считаем, что алфавит из n букв состоит из первых n натуральных чисел.

ОПРЕДЕЛЕНИЕ 1. Пусть дан алфавит Σ мощности n . Назовем *отображением Парика* функцию $\Phi : \Sigma^* \rightarrow \mathbb{N}^n$, которая сопоставляет каждому слову вектор, i -я компонента которого равна количеству вхождений i -й буквы в слово.

ОПРЕДЕЛЕНИЕ 2. *Линейным множеством* будем называть подмножество $\mathbb{N}^n$, состоящее из всех элементов вида $b + \sum_{i=1}^k \alpha_i a_i$, где b и a_i — фиксированные векторы из $\mathbb{N}^n$, а α_i — произвольные натуральные числа.

ОПРЕДЕЛЕНИЕ 3. *Полулинейным множеством* будем называть конечное объединение линейных множеств и пустое множество.

Теорема 1 [10, теорема 10.35]. Пусть дан контекстно-свободный язык L . Тогда образ Парика $\Phi(L)$ полулинеен.

Лемма 1 [10, теорема 2.27]. Язык L в однобуквенном алфавите регулярен тогда и только тогда, когда $\Phi(L)$ полулинеен. При этом линейные множества в $\mathbb{N}$ — это в точности арифметические прогрессии.

Лемма 2 [10, задача 9.24]. Пересечение контекстно-свободного языка и регулярного языка — контекстно-свободный язык.

ОПРЕДЕЛЕНИЕ 4. *Функцией роста языка* $L \subseteq \Sigma^*$ называется

$$\gamma(n) = |L \cap \Sigma^n|.$$

Теорема 2 [11, доказательство теоремы 1.1]. Пусть L — контекстно-свободный язык с функцией роста γ . Тогда верно одно из двух:

- (1) существуют слова $w_1, \dots, w_n$, такие, что $L \subseteq w_1^* \dots w_n^*$;

(2) существуют такие бесконечная арифметическая прогрессия P и числа $r > 1$, n_0 , что $\gamma(n) \geq r^n$ для всех $n \in P$, $n \geq n_0$.

ЗАМЕЧАНИЕ 1. В [11] сформулирован другой результат (теорема 1.1), который буквально неверен: в п. 2 утверждается, что $\gamma(n) \geq r^n$ для всех $n \geq n_0$. Доказательство теоремы 1.1 в [11] корректно для *кумулятивной функции роста* — количеству слов из языка, длины которых не превосходят n . Однако из этого доказательства следует и нужный нам факт — теорема 2. В [11] доказано, что если не выполняется п. 1 теоремы, то существуют такие слова x, y, u, v, w , для которых $|x| = |y|$ и все слова $W(x, y)wW^R(u, v)$ принадлежат L . Здесь $W(x, y)$ обозначает слово, которое получается из слова W в двухбуквенном алфавите подстановкой вместо первого символа x , а вместо второго — y ; W^R обозначает записанное в обратном порядке слово W ; а $W^R(u, v)$ обозначает слово, которое получается из W^R подстановкой вместо первого символа u , а вместо второго — v . Более того, для слов W одинаковой длины все слова вида $W(x, y)wW^R(u, v)$ попарно различны. Выбирая сбалансированные слова W длины $2k$ (с одинаковым количеством 0 и 1), получаем, что $\gamma(n) \geq \binom{2k}{k}$ для $n \in P = \{|w| + k(|x| + |y| + |u| + |v|) : k \in \mathbb{N}\}$, откуда и следует п. 2 теоремы 2.

Зачастую, например в [5–7, 12], при рассмотрении задач регулярной реализуемости входом считается описание некоторого НКА. Однако мы в дальнейшем будем считать, что на вход подается описание ДКА словом в двоичном алфавите. Точный вид такого описания неважен, требуется лишь, чтобы по описанию возможно было моделировать работу ДКА за полиномиальное время. Язык, распознаваемый автоматом A , будем обозначать через $\mathcal{L}(A)$, а описание автомата A — через $\langle A \rangle$. Далее будем считать, что все ДКА являются полными.

ОПРЕДЕЛЕНИЕ 5. Назовем *задачей регулярной реализуемости* для языка L следующий язык:

$$\text{DRR}_L = \{\langle A \rangle \mid |L \cap \mathcal{L}(A)| \neq 0\}.$$

ОПРЕДЕЛЕНИЕ 6. Назовем *задачей бесконечной регулярной реализуемости* для языка L следующий язык:

$$\text{DRR}_L^\infty = \{\langle A \rangle \mid |L \cap \mathcal{L}(A)| = \infty\}.$$

ОПРЕДЕЛЕНИЕ 7. Зададим морфизм $i(0) = 1, i(1) = 0$. Рассмотрим следующую последовательность слов в алфавите $\{0, 1\}$: $T_1 = 0, T_{k+1} = T_k i(T_k)$. Заметим, что каждое следующее слово включает предыдущее как префикс. Последовательностью *Туэ — Морса* будем называть такое бесконечное слово T , что первые его 2^k символов равны слову T_k . Будем обозначать первые n символов последовательности Туэ — Морса через $T(n)$.

Лемма 3 [13, гл. 2.5]. Не существует такого непустого слова w , что www является подсловом последовательности Туэ — Морса.

Лемма 4 [13, гл. 2.5]. На i -м месте в последовательности Туэ — Морса стоит 1 тогда и только тогда, когда количество единиц в двоичном представлении числа i нечетно (будем считать, что нумерация индексов позиций начинается с 0). Отсюда следует, что $T(n)$ вычислимо на логарифмической по n памяти.

Далее используем стандартное определение вычислимости на памяти $s(n)$ и классов $\text{SPACE}(s(n))$ (см., например, [14, определение 4.1]). Через $\text{Logspace} =$

$\text{SPACE}(\log n)$ обозначаем класс языков, разрешимых на логарифмической памяти.

ОПРЕДЕЛЕНИЕ 8. Сводимость языка M к языку L на логарифмической памяти — это m -сводимость, в которой сводящая функция вычислима на логарифмической памяти. Если M сводится к L на логарифмической памяти и L сводится к M на логарифмической памяти, то будем писать $M \equiv_l L$.

ОПРЕДЕЛЕНИЕ 9 [14, гл. 3.4]. Будем говорить, что функция *принадлежит классу* $\text{DTime}(f(n))^L$, если она вычислима за время $O(f(n))$ на детерминированной машине Тьюринга с оракулом L .

3. Булевы алгебры

Перечислим основные факты и определения теории булевых алгебр, которые понадобятся в ходе доказательств.

ОПРЕДЕЛЕНИЕ 10. Будем обозначать булеву алгебру регулярных языков над алфавитом из n букв через $\mathbf{REG}_n$. Если количество букв не играет роли, то нижний индекс будем опускать.

ОПРЕДЕЛЕНИЕ 11. Будем называть *относительно регулярной алгеброй* для языка L (и обозначать через $\mathbf{REG}_L$) булеву алгебру, образованную пересечениями регулярных языков с L . Операциями в ней являются объединение, пересечение и дополнение до L .

ОПРЕДЕЛЕНИЕ 12. Будем говорить, что элемент булевой алгебры a *меньше или равен* чем b и обозначать это через $a \leq b$, если существует элемент c такой, что $b \wedge c = a$.

ОПРЕДЕЛЕНИЕ 13. Будем называть ненулевой элемент a *атомом*, если верно следующее:

$$b \leq a \rightarrow (b = 0) \text{ или } (b = a).$$

ОПРЕДЕЛЕНИЕ 14. Алгебру A будем называть *безатомной*, если в ней нет атомов.

ОПРЕДЕЛЕНИЕ 15. Алгебру A будем называть *атомной*, если для любого ненулевого $a \in A$ существует такой атом $b \in A$, что $b \leq a$.

ОПРЕДЕЛЕНИЕ 16. Пусть A — булева алгебра. Множество $I \subset A$ называется *идеалом*, если I содержит 0, для каждого элемента i из I все меньшие i элементы также содержатся в I , а также I замкнуто относительно операции $\vee$. Идеал индуцирует отношение эквивалентности на булевой алгебре:

$$x \sim_I y \iff (x \wedge \neg y) \vee (y \wedge \neg x) \in I.$$

ОПРЕДЕЛЕНИЕ 17. Идеалом Фреше $F(A)$ булевой алгебры A называется идеал, состоящий из всех конечных дизъюнкций ее атомов (и нуля).

ОПРЕДЕЛЕНИЕ 18. Будем обозначать фактор булевой алгебры A по идеалу I через A/I . Отдельно введем обозначение $A' = A/F(A)$.

Лемма 5 [15, предложение 1.5.2]. Пусть A и B — счетные атомные булевы алгебры. Если A' изоморфна B' , то A изоморфна B .

Заметим, что относительно регулярные алгебры всегда являются счетными и атомными. Будем обозначать через $[R]_L$ класс эквивалентности языка R в

факторе $\mathbf{REG}_L$ по идеалу Фреше. Если L ясно из контекста, то нижний индекс будем опускать. При этом

$$[R]_L = [K]_L \iff |(R \cap L) \Delta (K \cap L)| < \infty.$$

Более того, справедливы следующие утверждения.

Теорема 3 [9, теорема 3]. *Для любой счетной атомной алгебры существует изоморфная ей относительно регулярная алгебра.*

Лемма 6 [16, предложение 12]. $\mathbf{REG}'$ является безатомной алгеброй.

Лемма 7 [15, предложение 1.5.1]. *Любые две безатомные счетные алгебры изоморфны.*

Лемма 8 [15, гл. 1.4]. *Если булева алгебра конечна, то в ней 2^n элементов для некоторого $n \in \mathbb{N}$. При этом для каждого натурального n существует единственная с точностью до изоморфизма булева алгебра с 2^n элементами.*

Лемма 9 [15, гл. 1.4]. *Рассмотрим подалгебру булевой алгебры, порожденную n элементами. Тогда количество элементов в ней не превышает 2^{2^n} . При этом если эти элементы ненулевые и не пересекаются, а в объединении дают единицу, то они являются ее атомами и тогда количество элементов в ней равно 2^n . Любой элемент подалгебры, порожденной конечным количеством элементов, является конечным объединением атомов.*

4. Примеры относительно регулярных алгебр

Лемма 10. *Если L — бесконечный язык, то $\mathbf{REG}_L \simeq \mathbf{REG}$ равносильно тому, что для любого регулярного языка R из $|R \cap L| = \infty$ следует существование такого регулярного языка K , что:*

- (1) $K \subset R$,
- (2) $|(R - K) \cap L| = \infty$,
- (3) $|K \cap L| = \infty$.

ДОКАЗАТЕЛЬСТВО. Фактор $\mathbf{REG}$ по идеалу Фреше — это безатомная алгебра по лемме 6. В силу леммы 5 и леммы 7 $\mathbf{REG}_L \simeq \mathbf{REG}$ равносильно тому, что $\mathbf{REG}'_L$ безатомная. Указанные в лемме условия проверяют, что для любого элемента в факторе по идеалу Фреше найдется меньший его ненулевой элемент. Действительно, рассмотрим произвольный регулярный R . Для того чтобы $[R]$ в $\mathbf{REG}'_L$ не оказался атомом, необходимо найти меньший его ненулевой элемент. Первое условие проверяет, что $[K]$ меньше или равен $[R]$, второе — что $[R] \neq [K]$, третье — что $[K] \neq [\emptyset]$.

Теорема 4. *Пусть $L \subseteq \Sigma^*$ — такой бесконечный язык, что для любого регулярного языка R из $|R \cap L| = \infty$ следует, что $\Phi(L \cap R)$ содержит бесконечное линейное множество. Тогда $\mathbf{REG}_L \simeq \mathbf{REG}$.*

ДОКАЗАТЕЛЬСТВО. В доказательстве будем пользоваться леммой 10. Пусть $|R \cap L| = \infty$ для регулярного языка R . Тогда по условию теоремы $P \subseteq \Phi(L \cap R)$, где $P = b + \sum_{i=1}^k \alpha_i a_i$ — бесконечное линейное множество. Хотя бы одна из координат хотя бы одного вектора a_i строго положительна. Пусть $a_j^s > 0$ (номера координат в верхнем индексе). Положим $W = \{w \mid \Phi(w)^s \equiv b^s \pmod{2a_j^s}\}$. Это

такие слова, у которых s -я координата вектора Парика совпадает с s -й координатой вектора $b + 2\alpha a_j$, $\alpha \in \mathbb{N}$. Обозначим $Q = R \cap W$. Этот язык регулярен как пересечение двух регулярных языков.

Заметим, что Q удовлетворяет всем трем условиям леммы 10:

- (1) $Q \subset R$;
- (2) $Q \cap L$ содержит бесконечное множество слов, векторы Парика которых имеют вид $b + 2\alpha a_j$, $\alpha \in \mathbb{N}$;
- (3) $(R - Q) \cap L$ содержит бесконечное множество слов, векторы Парика которых имеют вид $b + (2\alpha + 1)a_j$, $\alpha \in \mathbb{N}$.

ЗАМЕЧАНИЕ 2. Условия теоремы 4 выполняются, если $\Phi(L \cap R)$ полулинейное для любого регулярного языка R . В частности, из теорем 1, 4 и леммы 2 следует, что для любого бесконечного контекстно-свободного языка L выполняется $\mathbf{REG}_L \simeq \mathbf{REG}$.

Лемма 11. Пусть $L = \{0^n 1^n 2^n \mid n \in \mathbb{N}\}$. Язык L не является контекстно-свободным, но $L \in \mathbf{Logspace}$. При этом $\mathbf{REG}_L \simeq \mathbf{REG}$.

ДОКАЗАТЕЛЬСТВО. Докажем, что $\mathbf{REG}_L \simeq \mathbf{REG}$. Для этого воспользуемся теоремой 4. Покажем, что язык $U = \{0^n \mid 0^n 1^n 2^n \in R\}$ является регулярным для любого регулярного R . Из этого следует, что $\Phi(L \cap R)$ полулинейно.

Итак, пусть язык R распознается ДКА $A = (Q, s, F, \delta)$. Построим НКА $\hat{A} = (\hat{Q}, \hat{s}, \hat{F}, \hat{\delta})$ для языка U :

- $\hat{Q} = \{\hat{s}\} + Q^5$;
- $\hat{F} = \{(q_1, q_1, q_2, q_2, f) \mid q_1, q_2 \in Q, f \in F\}$;
- $\hat{\delta}(\hat{s}, \varepsilon) = \{(s, q_1, q_1, q_2, q_2) \mid q_1, q_2 \in Q\}$;
- $\hat{\delta}((x, q_1, y, q_2, z), 0) = \{(\delta(x, 0), q_1, \delta(y, 1), q_2, \delta(z, 2))\}$;

$\hat{A}$ принимает слово 0^n тогда и только тогда, когда

$$\delta(s, 0^n) = q_1, \quad \delta(q_1, 1^n) = q_2, \quad \delta(q_2, 2^n) = q_f, \quad q_f \in F.$$

Это равносильно тому, что $0^n 1^n 2^n \in R$.

Лемма 12. Пусть $L = \{0^{n!} \mid n \in \mathbb{N}\}$. Язык L не является контекстно-свободным, но $L \in \mathbf{Logspace}$. При этом $\mathbf{REG}_L \not\simeq \mathbf{REG}$.

ДОКАЗАТЕЛЬСТВО. Покажем, что фактор $\mathbf{REG}_L$ по идеалу Фреше состоит из двух элементов. Рассмотрим произвольный бесконечный унарный регулярный язык R . Он является конечным объединением арифметических прогрессий. Рассмотрим конкретную прогрессию $P = \{kn + b \mid n \in \mathbb{N}\}$. Если $k = 0$, то она конечна. Если $b \equiv 0 \pmod{k}$, то P содержит все кратные k за исключением конечного множества. Поскольку $m! \equiv 0 \pmod{k}$ при $m \geq k$, то в данном случае P включает в себя все факториалы за исключением конечного множества. Если $b \not\equiv 0 \pmod{k}$, то все элементы P не равны нулю по модулю k . Значит, в P попадает конечное количество факториалов.

Мы получили, что либо $\{n! \mid n \in \mathbb{N}\} \setminus P$ конечно, либо $\{n! \mid n \in \mathbb{N}\} \cap P$ конечно. Значит, и для произвольного регулярного языка R в унарном алфавите это верно. Получаем, что фактор $\mathbf{REG}_L$ по идеалу Фреше состоит из двух элементов.

Теорема 5. Рассмотрим произвольный нетривиальный язык L в алфавите из двух букв, т. е. $L \neq \emptyset$ и $\bar{L} \neq \emptyset$. Тогда существует такой язык $\hat{L}$, что $\hat{L} \equiv_l L$, и при этом $\mathbf{REG}_{\hat{L}} \simeq \mathbf{REG}$.

ДОКАЗАТЕЛЬСТВО. Используем следующие обозначения:

- i — морфизм из определения 7 (т. е. $i(0) = 1, i(1) = 0$);
- $\psi(w) = wT(|w|)0^{|w|^2-2|w|}$ при $|w| > 2$, $\psi(w) = w$ при $|w| \leq 2$;
- $h(w) = \psi(w)i(\psi(w))$;
- Σ^n — множество всех двоичных слов длины n ;
- P^n — множество двоичных слов длины n , в которых одинаково количество нулей и единиц;
- B^n — пересечение P^n с объединением всех языков вида $w_1^*w_2^*\dots w_t^*$, где $\sum_{i=1}^t |w_i| \leq \log n$;
- $H^n = \{h(w) \mid |h(w)| = n\}$;
- $H_L^n = \{h(w) \mid w \notin L, |h(w)| = n\}$;
- $\hat{L}^n = P^n - (B^n \cup H_L^n)$;
- для каждого из введенных классов языков A^n обозначим $A = \bigcup_{n=0}^{\infty} A^n$;

и в другую сторону: для языка A обозначаем $A^n = \{w \in A \mid |w| = n\}$.

Заметим, что H^n и B^n вложены в P^n и, начиная с некоторого n , множества H^n и B^n не пересекаются. Действительно, предположим противное: $h(w) \in w_1^*w_2^*\dots w_t^*$, где $\sum_{i=1}^t |w_i| \leq \log n$. Заметим, что $h(w)$ содержит начало

последовательности Туэ — Морса длины $|w| = \sqrt{n/2}$. Если пересечение под слова $T(|w|)$ с некоторым блоком w_i^* длиннее $3|w_i|$, то внутри последовательности Туэ — Морса возникает куб некоторого циклического сдвига слова w_i , что противоречит лемме 3. Отсюда получаем неравенство

$$\sqrt{n/2} \leq 3 \sum_{i=1}^t |w_i| \leq 3 \log n,$$

которое не выполняется при достаточно больших n .

Для начала докажем, что $\mathbf{REG}_{\hat{L}} \simeq \mathbf{REG}$. В доказательстве будем пользоваться теоремой 4. Так как $\hat{L}^n \subset P^n$, то $\Phi(\hat{L}) \subseteq \{\alpha \overrightarrow{(1,1)} \mid \alpha \in \mathbb{N}\}$. Рассмотрим некоторый регулярный R . Пусть $\tilde{R} = P \cap R$. Язык $\tilde{R}$ контекстно-свободный, поэтому по теореме 2 либо существуют такие слова $w_1, \dots, w_t$, что $\tilde{R} \subset w_1^* \dots w_t^*$, и тогда $\hat{L} \cap R$ конечно, так как $\tilde{R}^n \subset B^n$, начиная с некоторого n ; либо функция роста языка $\tilde{R}$ экспоненциальна на некоторой бесконечной арифметической прогрессии длин E , состоящей только из четных длин (исходя из определения, P^n непусто только для четных n). В последнем случае покажем, что $\hat{L}^n \cap \tilde{R}^n \neq \emptyset$ при достаточно больших $n \in E$. Отсюда следует, что

$$\Phi(\hat{L} \cap R) = \Phi(\hat{L} \cap \tilde{R}) \supseteq \{(n/2, n/2) \mid n \in E, n \geq n_0\}$$

и $\mathbf{REG}_{\hat{L}} \simeq \mathbf{REG}$ по теореме 4.

Оценим размеры описанных выше множеств при достаточно больших n . Всего слов в P^n столько, сколько есть способов разместить $\frac{n}{2}$ единиц в слове длины n , т. е. $\binom{n}{\frac{n}{2}} = 2^{n-o(n)}$. Языков вида $w_1^*w_2^*\dots w_t^*$, где $\sum_{i=1}^t |w_i| \leq \log n$, не больше, чем количество разбиений слов длины $\log n$ на непустые под слова, что не больше $2^{\log n} 2^{\log n} = n^2$. Количество слов каждого такого языка в B^n не больше $n^t \leq n^{\log n}$. Поэтому $|B^n| \leq n^2 2^{\log^2 n} = 2^{o(n)}$. В H^n слов не больше, чем

различных прообразов h , т. е. $2^{\sqrt{\frac{n}{2}}} = 2^{o(n)}$. Получаем, что в $P^n - \hat{L}^n \subseteq B^n \cup H^n$ количество слов равно $2^{o(n)}$. С другой стороны, для достаточно больших $n \in E$ количество слов длины n в языке $\tilde{R}$ не меньше $2^{\rho n}$, $\rho > 0$. Поэтому $\hat{L}^n \cap \tilde{R}^n \neq \emptyset$ при достаточно больших $n \in E$. Как объяснялось выше, отсюда следует применимость теоремы 4.

Теперь построим указанные сводимости. Введенная ранее функция h задает сводимость L к $\hat{L}$ на словах достаточно большой длины. Для остального (конечного) множества слов сводящую функцию определяем произвольно с соблюдением условия сводимости, это не влияет на оценку вычислительной сложности сводящей функции. Арифметические операции, как и последовательность Туэ — Морса (см. лемму 4), вычислимы на логарифмической памяти, поэтому h также вычислима на логарифмической памяти. Обратная сводимость устроена сложнее. Будем строить функцию f , сводящую $\hat{L}$ к L . Пусть $Y \in L$ и $N \notin L$ — два фиксированных слова. Определим $f(x)$ следующим образом.

- Если $x \notin P$ (что проверяется на логарифмической памяти), то $f(x) = N$.
- Если $x \in B$, то $f(x) = N$.

Принадлежность B равносильна тому, что $x \in B^{|x|}$. Будем последовательно проверять, имеет ли x вид $w_1^* \dots w_t^*$ для конкретных $w_1, \dots, w_t$. Как было указано раньше, количество различных наборов w_i для $B^{|x|}$ полиномиально по $|x|$. Тогда для хранения номера проверяемого в конкретный момент набора слов достаточно логарифмической по $|x|$ памяти. По конкретному набору слов можно построить НКА, проверяющий слово на принадлежность $w_1^* \dots w_t^*$ со следующими свойствами: количество состояний в нем линейно по суммарной длине слов w_i и количество переходов из каждого состояния не более двух. Таким образом, для хранения данного НКА достаточно логарифмической памяти. Принятие слова таким НКА проверяется следующим образом. Нужно хранить список состояний НКА, достижимых к данному шагу работы, и последовательно обновлять его при чтении входа. Хранить данный список можно в виде битовой маски: двоичного слова длины, равной количеству состояний, в котором 1 отмечает те состояния, которые достижимы из начального за k шагов. Обновление такого списка возможно на логарифмической памяти. Последовательно читая символы x , поддерживаем множество достижимых состояний НКА; после обработки всех символов проверяем, содержит ли это множество принимающее состояние.

Если $x \in H$, то $f(x)$ — это первые $\sqrt{\frac{n}{2}}$ символов x для достаточно длинных слов. На конечном множестве слов малой длины сводящая функция определяется произвольно с соблюдением условия сводимости.

Если $x \in P - B - H$, то $f(x) = Y$.

Следствие 1. Существует неразрешимый (и даже не арифметичный) язык L , для которого $\mathbf{REG}_L \simeq \mathbf{REG}$.

5. Вычислимость изоморфизмов относительно регулярных алгебр

Будем считать, что каждый элемент $P \in \mathbf{REG}_L$ задается (не единственным образом) таким автоматом M , что $\mathcal{L}(M) \cap L = P$. Таким образом, для перечисления всех элементов алгебры можно просто перечислять все автоматы. Аналогичным образом задаются и факторы по идеалу Фреше для данных ал-

гебр: так как элементы данной алгебры — это классы эквивалентности языков, то каждый элемент Q задается таким автоматом M , что $\mathcal{L}(M) \cap L$ лежит в классе Q . Доказательства следующих фактов, по сути, являются конструктивизацией частных случаев критерия Вoota (см. [15]).

Теорема 6. Пусть $\mathbf{REG}_L \simeq \mathbf{REG}$ и $\mathbf{DRR}_L^\infty$ разрешим. Тогда существует вычислимый изоморфизм f алгебр $\mathbf{REG}'_L$ и $\mathbf{REG}'$.

ДОКАЗАТЕЛЬСТВО. Для начала зафиксируем такие перечисления (с повторениями) $A_1, A_2, \dots$ элементов $\mathbf{REG}'_L$ и $B_1, B_2, \dots$ элементов $\mathbf{REG}'$, в которых все ДКА перечисляются по возрастанию количества состояний в них.

Построение f проводится по индукции, так что каждый раз его область определения расширяется и является некоторой конечной подалгеброй. При этом должно выполняться условие: на шаге индукции под номером $2t + 1$ подалгебра включает в себя элемент A_t , а на шаге индукции под номером $2t + 2$ образ подалгебры включает в себя элемент B_t . Таким образом, получив автомат из n состояний на вход, необходимое количество шагов индукции не превышает количество автоматов, которые имеют не более n состояний.

После каждого шага построенный изоморфизм f_t хранится следующим образом: атомы подалгебры в $\mathbf{REG}'_L$ и значения f_t на них.

База индукции: для начала фиксируем $f(\emptyset) = \emptyset$, $f(L) = \Sigma^*$.

Нечетный шаг индукции ($2t + 1$): пусть мы уже имеем некоторую конечную подалгебру в $\mathbf{REG}'_L$ и знаем набор ее атомов $a_1, \dots, a_l$, которые задаются автоматами $M_1, \dots, M_l$. Построим элементы $\hat{a}_{2i} = a_i \cap A_t$, $\hat{a}_{2i+1} = a_i \cap \overline{A_t}$. Это кандидаты в атомы расширенной подалгебры. Такое построение вычислимо, так как по двум автоматам P_1 и P_2 можно алгоритмически построить автомат P_3 , такой, что $\mathcal{L}(P_3) = \mathcal{L}(P_1) \cap \mathcal{L}(P_2)$.

Заметим, что получившиеся элементы попарно не пересекаются, за исключением двух следующих случаев.

1. Некоторые из них могут задавать тот же элемент в алгебре, что и пустое множество, для них образ уже определен. Это эквивалентно конечности пересечения с L , что можно проверить с помощью оракула бесконечной регулярной реализуемости.

2. Некоторые из них могут задавать тот же элемент, что и исходный a_i , для них образ также уже определен. Равносильно, симметрическая разность получившегося элемента и a_i имеет конечное пересечение с L , что также можно проверить с помощью оракула бесконечной регулярной реализуемости.

Для остальных элементов найдем такой $\hat{b}_{2i}$, что $0 < \hat{b}_{2i} < f_t(a_i)$, и доопределим $\hat{b}_{2i+1} = f_t(a_i) - \hat{b}_{2i}$. Укажем конкретную процедуру поиска такого элемента. Так как оба фактора безатомные, то искомым элемент всегда существует. Для его построения в образе $f_t(a_i)$ найдем такой бесконечный регулярный подязык, что его дополнение до $f_t(a_i)$ тоже бесконечно. Для этого можно, например, найти, какие прогрессии образуют длины слов в $f_t(a_i)$ и взять пересечение $f_t(a_i)$ с языком слов, длины которых принадлежат собственной бесконечной подпрогрессии. После этого мы задаем образы так: $f(\hat{a}_i) = \hat{b}_i$. Шаг индукции завершен.

Четный шаг индукции ($2t + 2$): в этот раз построим элементы $\hat{b}_{2i} = b_i \cap B_t$, $\hat{b}_{2i+1} = b_i \cap \overline{B_t}$, аналогично проверив на совпадение нового элемента с исходным или с нулем алгебры. В данном случае для этого достаточно проверить,

что автоматы, представляющие $\widehat{b}_{2i}$ и $b_i - \widehat{b}_{2i}$, задают бесконечные языки. Если совпадения нет, ищем такой $\widehat{a}_{2i}$, что $0 < \widehat{a}_{2i} < f_t^{-1}(b_i) = a_i$, и доопределяем $\widehat{a}_{2i+1} = f_t^{-1}(b_i) - \widehat{a}_{2i}$. Построение такого набора вычислимо. Действительно, пусть элемент a_i имеет автомат-представителя M_i . Подходящий элемент точно существует, так как $\mathbf{REG}'_L$ безатомная. Поэтому его можно найти, перебирая все автоматы и проверяя для каждого автомата P следующие условия: $\mathcal{L}(P) \cap \mathcal{L}(M_i) \cap L$ бесконечен и $\overline{\mathcal{L}(P)} \cap \mathcal{L}(M_i) \cap L$ бесконечен. Это возможно с использованием оракула бесконечной регулярной реализуемости. После этого полагаем $f(\widehat{a}_{2i}) = \widehat{b}_{2i}$, $f(\widehat{a}_{2i+1}) = \widehat{b}_{2i+1}$.

Так как графики функций f_t — расширяющиеся множества, и за два шага индукции туда попадает в прообраз очередной A_t , а в образ B_t , то объединение f_t и даст искомый изоморфизм.

Отметим, что так как на четных шагах определяется прообраз очередного элемента из $\mathbf{REG}'$, то обратное отображение также вычислимо.

Теорема 7. Пусть $\mathbf{REG}_L \simeq \mathbf{REG}$ для некоторого языка $L \subset \Sigma^*$, а $\mathbf{DRR}_L$ и $\mathbf{DRR}_L^\infty$ разрешимы. Тогда существует вычислимый изоморфизм $f : \mathbf{REG}_L \rightarrow \mathbf{REG}$.

ДОКАЗАТЕЛЬСТВО. Общая идея данного доказательства совпадает с доказательством теоремы 6. При построении изоморфизма f будем считать, что по теореме 6 уже построен изоморфизм g факторов по идеалам Фреше. Изоморфизм f поднимает изоморфизм g на алгебру, т. е. $[f(x)] = g([x])$.

Определим функцию $\mathbf{DRR}_L^C(R)$ на регулярных языках как $|L \cap R|$, если пересечение $L \cap R$ конечно, и как ∞ , если пересечение бесконечно. Эта функция вычислима с оракулами $\mathbf{DRR}_L$ и $\mathbf{DRR}_L^\infty$. Для начала вычислим $\mathbf{DRR}_L^\infty(R)$, если $L \cap R$ бесконечно, то $\mathbf{DRR}_L^C(R) = \infty$. Затем вычислим $\mathbf{DRR}_L(R)$, если $L \cap R = \emptyset$, то $\mathbf{DRR}_L^C(R) = 0$. Иначе $0 < \mathbf{DRR}_L^C(R) < \infty$. Будем проверять все слова по порядку, постепенно строя пересечение L с R . Для проверки $w \in L \cap R$ проверяем $w \in R$ обычным алгоритмом проверки принадлежности слова регулярному языку и проверяем $w \in L$ вызовом оракула $\mathbf{DRR}_L(\{w\})$. После обнаружения очередного слова из $L \cap R$ проверяем совпадение $L \cap R$ и (конечного) языка W уже найденных слов из $R \cap L$. Для этого вычисляем $\mathbf{DRR}_L(R - W)$. Если $L \cap (R - W) = \emptyset$, то $W = R \cap L$ и $\mathbf{DRR}_L^C(R) = |W|$.

Зафиксируем такие перечисления $A_1, A_2, A_3, A_4, \dots$ элементов $\mathbf{REG}_L$ и $B_1, B_2, B_3, B_4, \dots$ элементов $\mathbf{REG}$, в которых все ДКА перечисляются по возрастанию количества состояний в них.

Построение функции f будет происходить по индукции, аналогичной индукции в теореме 6.

Опишем нечетный шаг индукции. Пусть уже построена конечная подалгебра в $\mathbf{REG}_L$ и известен набор ее атомов $a_1, \dots, a_n$. При этом должны быть выполнены условия (инварианты цикла) $\mathbf{DRR}_L^C(a_i) = \mathbf{DRR}_\Sigma^C(b_i)$, где $b_i = f(a_i)$, и $[b_i] = [g(a_i)]$ в $\mathbf{REG}'$. Заметим, что для базы индукции ($f(\emptyset) = \emptyset, f(\Sigma^*) = \Sigma^*$) эти условия выполняются. Построим элементы $\widehat{a}_{2i} = a_i \cap A_t, \widehat{a}_{2i+1} = a_i \cap \overline{A_t}$. Это кандидаты в атомы расширенной подалгебры. Так как мы считаем, что оба элемента задаются автоматами, то такое построение вычислимо.

Заметим, что получившиеся элементы попарно не пересекаются, за исключением двух следующих случаев.

1. Некоторые из них могут задавать тот же элемент в алгебре, что и пустое

множество, для них образ уже определен. Это эквивалентно пустоте пересечения с L , что возможно проверить с помощью оракула регулярной реализуемости.

2. Некоторые из них могут задавать тот же элемент, что и исходный a_i , для них образ тоже уже определен. Равносильно, симметрическая разность получившегося элемента и a_i не пересекается с L , что возможно проверить с помощью оракула регулярной реализуемости.

Для остальных i укажем конкретную процедуру поиска $\hat{b}_{2i}$. Для начала вычислим $g(\hat{a}_{2i})$, $\text{DRR}_L^C(\hat{a}_{2i})$ и $\text{DRR}_L^C(a_i - \hat{a}_{2i})$. Рассмотрим несколько случаев:

1) $\text{DRR}_L^C(\hat{a}_{2i}) = m < \infty$, тогда в качестве $\hat{b}_{2i}$ выберем дизъюнкцию m атомов, лежащих под $f(a_i)$, что возможно в силу инварианта цикла, выполняющегося на предыдущем шаге;

2) $\text{DRR}_L^C(a_i - \hat{a}_{2i}) = m < \infty$, тогда в качестве $\hat{b}_{2i}$ выберем дополнение в $f(a_i)$ к дизъюнкции m атомов, лежащих под $f(a_i)$, что возможно в силу инварианта цикла, выполняющегося на предыдущем шаге;

3) иначе в качестве $\hat{b}_{2i}$ выберем $g(\hat{a}_{2i}) \wedge f(a_i)$.

В каждом случае инварианты цикла сохраняются. В частности, в третьем случае это следует из равенств

$$[\hat{b}_{2i}] = [g(\hat{a}_{2i}) \wedge f(a_i)] = g(\hat{a}_{2i}) \wedge g(a_i) = g(\hat{a}_{2i} \wedge a_i) = g(\hat{a}_{2i}).$$

В первых двух случаях сохранение инварианта прямо следует из правила выбора $\hat{b}_{2i}$.

Опишем, как можно вычислимо выбрать соответствующие подмножества атомов. Пусть $\hat{a}_{2i}$ имеет конечное пересечение с L . Надо выбрать m различных слов в языке (детерминированного) автомата M_i , задающего $b_i = f(a_i)$. Для этого достаточно построить (например, обходом в ширину) m различных путей из стартового состояния в финальное в графе этого автомата. Таким путям соответствуют слова из $\mathcal{L}(M_i)$.

После выбора $\hat{b}_{2i}$ полагаем $\hat{b}_{2i+1} = f(a_i) - \hat{b}_{2i} = b_i - \hat{b}_{2i}$.

На четном шаге индукции процедура аналогична, отдельно подчеркнем вычислимость построения. Необходимо найти m слов в пересечении некоторого регулярного языка и L . Это, конечно, вычислимо с оракулом DRR_L : достаточно перебирать слова из регулярного языка и проверять вхождение в L вызовом оракула $\text{DRR}_L(\{w\})$.

Отдельный интерес представляет исследование изоморфизмов между булевыми алгебрами регулярных языков над алфавитами разной мощности. Как можно заметить, предыдущие результаты доказывают существование вычислимого изоморфизма между алгебрами регулярных языков в алфавитах из двух букв и из одной буквы. Теперь дадим нижнюю оценку на сложность данного изоморфизма.

Теорема 8. Любой вычислимый изоморфизм из REG_2 в REG_1 имеет сложность $\Omega(2^{\frac{n}{2}})$, где n — количество состояний автомата, поданного на вход.

ДОКАЗАТЕЛЬСТВО. Рассмотрим слова из двух букв длины $m \geq 2$ в алфавите $\{0, 1\}$. Всего таких слов 2^m . По каждому слову w построим следующий автомат A_w :

- $Q = \{q_s\} \cup \{x \in \mathbb{N} \mid x < m\} \cup \{T\}$;

- $q_s \notin \{0, \dots, m-1, R\}$;
- $F = \{0\}$;
- $\delta(q_s, w_0) = 1$;
- $\delta(x, a) = x + 1 \bmod m$ если $w[x] = a$;
- $\delta(T, 0) = \delta(T, 1) = T$;
- все остальные переходы ведут в состояние T .

Легко видеть, что $\mathcal{L}(A_w) = w^+ = ww^*$. Заметим, что языки таких автоматов не пересекаются. Действительно, пусть $v, w \in \{0, 1\}^m$ и $u \in v^+ \cap w^+$. Тогда $u = v^k = w^k$, $k > 0$, а значит, $v = w$. Породим языками данных автоматов и дополнением к объединению всех таких языков подалгебру C в $\mathbf{REG}_2$. По лемме 9 в ней $2^m + 1$ атомов и 2^{2^m+1} элементов. При этом все атомы кроме одного задаются автоматами с $m + 2$ состояниями.

Теперь оценим количество непересекающихся унарных регулярных языков, задаваемых автоматами с не более чем n состояниями. Их не больше, чем непересекающихся арифметических прогрессий с разностью, не превосходящей n , а таких не более чем $O(n^2)$.

Любой изоморфизм должен переводить подалгебру в подалгебру, а атомы подалгебры — в атомы ее образа. Рассмотрим атомы образа C , отвечающие языкам $\mathcal{L}(A_w)$. Их 2^m штук, они не пересекаются, но тогда количество состояний, которыми задается хотя бы один из них $\Omega(2^{\frac{n}{2}})$, где $n = m + 2$ — количество состояний автомата A_w . Заметим, что время работы алгоритма вычисления изоморфизма не меньше длины выдаваемого описания автомата, а длина стандартного описания автомата оценивается снизу числом его состояний. Отсюда получаем нижнюю оценку на время работы алгоритма вычисления изоморфизма, указанную в теореме.

Благодарности. Авторы благодарны рецензенту за полезные замечания и предложения, существенно улучшившие содержание работы.

ЛИТЕРАТУРА

1. Vyalyi M. On models of a nondeterministic computation // Lecture Notes in Computer Science. 2009. V. 5675. P. 334–345.
2. Bouajjani A., Esparza J., Maler O. Reachability analysis of pushdown automata: Application to model-checking // Lecture Notes in Computer Science. 1997. V. 1243. P. 135–150.
3. Chistikov D., Majumdar R., Schepper P. Subcubic certificates for CFL reachability // Proc. ACM Program. Lang. 2022. V. 6. Article 41. 29 pp.
4. Koutiris P., Deep S. The fine-grained complexity of CFL reachability // Proc. ACM Program. Lang. 2023. V. 7. Article 59. 27 pp.
5. Wolf P. From decidability to undecidability by considering regular sets of instances // Theoretical Computer Sci. 2022. V. 899. P. 25–38.
6. Wolf P., Fernau H. Regular Intersection emptiness of graph problems: Finding a needle in a haystack of graphs with the help of automata // 2020. V. 29. <https://arxiv.org/abs/2003.05826>.
7. Diekert V., Fernau H., Wolf P. Properties of graphs specified by a regular language // Acta Informatica. 2022. V. 59. P. 357–385.
8. Wolf P. On the decidability of finding a positive ILP-instance in a regular set of ILP-instances // Acta Informatica. 2022. V. 59. P. 505–519.
9. Шиманогов И. Н., Вялый М. Н. Классификация относительно регулярных алгебр // Тр. МФТИ. 2024. Т. 16, № 4. С. 128–134.
10. Голубенко Д. А., Саватеев Ю. В. Языки, автоматы и грамматики. М.: МЦНМО, 2023.
11. Bridson M. R., Gilman R. H. Context-free languages of sub-exponential growth // J. Computer System Sci. 2002. V. 64, N 2. P. 308–310.
12. Anderson T., Loftus J., Rampersad N., Santean N., Shallit J. Detecting palindromes, patterns and borders in regular languages // Information and Computation. 2009. V. 207. P. 1096–1118.

13. Shallit J. A second course in formal languages and automata theory. Cambridge: Camb. Univ. Press, 2008.
14. Arora S., Barak B. Computational complexity: A modern approach. Cambridge: Camb. Univ. Press, 2009.
15. Гончаров С. С. Счетные булевы алгебры и разрешимость. Новосибирск: Науч. книга, 1996.
16. Selivanov V., Kononov A. Boolean algebras of regular languages // Developments in language theory. Berlin; Heidelberg: Springer, 2011. P. 386–396. (Lecture Notes in Computer Sci.; V. 6795)

Поступила в редакцию 10 декабря 2025 г.

После доработки 18 июня 2026 г.

Принята к публикации 24 июня 2026 г.

Шиманогов Игорь Николаевич (ORCID 0009-0003-9450-7119)

Московский физико-технический институт
(национальный исследовательский университет),
ул. Керченская, 1А, корп. 1, Москва 117303
shimanogov.in@phystech.edu

Вялый Михаил Николаевич (ORCID 0000-0001-9822-1060)

Московский физико-технический институт
(национальный исследовательский университет),
ул. Керченская, 1А, корп. 1, Москва 117303;
Федеральный исследовательский центр
«Информатика и управление» Российской академии наук,
ул. Вавилова, 44, корп. 2, Москва 119333;
Национальный исследовательский университет «Высшая школа экономики»,
Покровский б-р., 11, Москва 109028
vyalyi@gmail.com