Контекстно-свободные языки

Сборник задач

А. Е. Пентус, М. Р. Пентус

Москва 2016

Пентус А. Е., Пентус М. Р.
П25 Контекстно-свободные языки. Сборник задач. М.: РГГУ, 2016. 75 с.

Пособие содержит более 400 задач для освоения одного из классических разделов теории формальных языков. Рассматриваются контекстно-свободные грамматики и автоматы с магазинной памятью.

Для студентов и преподавателей математических, компьютерных и лингвистических направлений.

УДК 519.713
ББК 28.25

Рецензент —к. ф.-м. н. В. Н. Крупский

ISBN 978-5-7281-2012-4

Оглавление

Предисловие

Цель этого задачника — обеспечить различные курсы математической лингвистики задачами по контекстно-свободным языкам. Контекстно-свободные языки задаются двумя способами: контекстно-свободной грамматикой и автоматом с магазинной памятью.

Разбиение задач на разделы в основном соответствует учебнику «Математическая теория формальных языков». Чтобы задачник можно было использовать и без учебника, в каждом подразделе даются определения и теоремы (иногда с доказательством), необходимые для понимания и решения задач.

В первом разделе задачника рассматриваются самые общие понятия теории формальных языков и грамматик, определяется класс контекстно-свободных грамматик и формулируются необходимые теоремы об автоматных языках. В раздее 2 вводятся однозначные контекстно-свободные грамматики, в разделе 3 — нормальные формы Хомского и Грейбах. В задачах раздела 4 используются леммы о разрастании и замкнутость класса контекстно-свободных языков относительно ряда операций. Раздел 5 посвящён автоматам с магазинной памятью, раздел 6 — гомоморфизмам, раздел 7 — детерминированным автоматам с магазинной памятью.

Изложение строго математическое, но используются только самые простые математические понятия. Задачник можно рекомендовать для преподавания теории контекстно-свободных языков студентам направлений «Математика», «Прикладная математика и информатика», «Фундаментальная информатика и информационные технологии», «Информатика и вычислительная техника», «Лингвистика», «Фундаментальная и прикладная лингвистика», «Интеллектуальные системы в гуманитарной сфере». Авторы использовали часть задач из данного задачника в рамках курсов «Математическая лингвистика», «Математическая теория грамматик» и «Теория формальных языков». Задачи могут быть полезны также на практических занятиях по курсам «Теоретические основы информатики», «Теория автоматов и формальных языков», «Теория вычислительных процессов и структур».

1. Слова, языки и грамматики

1.1. Формальные языки

[20, 1.1], [6, с. 12—14], [2, 0.2.1, 0.2.2], [23, 1.1], [7, 1.1], [27, 1.5.1—1.5.3], [18, I.1.1], [9, с. 347—349], [24, 2.1], [25, с. 11—12], [5, с. 3—4], [3, 9.1], [19, 1.1], [11, 1.1, 1.3.1], [31, 1.7], [12, с. 152—153], [22, с. 22—23], [13, 1.3.1], [14, с. 257—262], [10, 2.2], [1, 3.3], [29, 1,4], [17, 2.6, 5.2, 6.2]

Определение 1. Будем называть *натуральными числами* неотрицательные целые числа. Множество всех натуральных чисел $\{0, 1, 2, \ldots\}$ обозначается $\mathbb{N}$.

Определение 2. *Алфавитом* называется конечное непустое множество. Его элементы называются *символами* (*буквами*).

Определение 3. *Словом* (*цепочкой*, *строкой*, string) в алфавите Σ называется конечная последовательность элементов Σ.

Пример 4. Рассмотрим алфавит $\Sigma = \{a, b, c\}$. Тогда $baaa$ является словом в алфавите Σ.

Определение 5. Слово, не содержащее ни одного символа (то есть последовательность длины 0), называется *пустым словом* и обозначается ε.

Определение 6. Множество всех слов в алфавите Σ обозначается Σ^*.

Определение 7. Множество всех непустых слов в алфавите Σ обозначается Σ^+.

Пример 8. Если $\Sigma = \{a\}$, то $\Sigma^+ = \{a, aa, aaa, aaaa, \ldots\}$.

Определение 9. Если $L \subseteq \Sigma^*$, то L называется *языком* (или *формальным языком*) над алфавитом Σ.

Поскольку каждый язык является множеством, можно рассматривать операции объединения, пересечения и разности языков, заданных над одним и тем же алфавитом (обозначения $L_1 \cup L_2$, $L_1 \cap L_2$, $L_1 - L_2$).

Пример 10. Множество $\{a, abb\}$ является языком над алфавитом $\{a, b\}$.

Определение 11. Пусть $L \subseteq \Sigma^*$. Тогда язык $\Sigma^* - L$ называется *дополнением* (complement) языка L относительно алфавита Σ. Когда из контекста ясно, о каком алфавите идёт речь, говорят просто, что язык $\Sigma^* - L$ является дополнением языка L.

Определение 12. Если x и y — слова в алфавите Σ, то слово xy (результат приписывания слова y в конец слова x) называется *конкатенацией* (*катенацией*, *сцеплением*) слов x и y. Иногда конкатенацию слов x и y обозначают $x \cdot y$.

Определение 13. Если x — слово и $n \in \mathbb{N}$, то через x^n обозначается слово $\underbrace{xx\ldots x}_{n \text{ раз}}$. Положим $x^0 \leftrightharpoons \varepsilon$ (знак $\leftrightharpoons$ читается «равно по определению»). Всюду далее показатели над словами и символами, как правило, являются натуральными числами.

Пример 14. По принятым соглашениям $ba^3 = baaa$ и $(ba)^3 = bababa$.

Пример 15. Множество $\{a^k ba^l \mid k \leqslant l\}$ является языком над алфавитом $\{a, b\}$. Этот язык содержит слова b, ba, aba, baa, $abaa$, $baaa$, $aabaa$, $abaaa$, $baaaa$ и т. д.

Определение 16. *Длина* слова w, обозначаемая $|w|$, есть число символов в w, причём каждый символ считается столько раз, сколько раз он встречается в w.

Пример 17. Очевидно, что $|baaa| = 4$ и $|\varepsilon| = 0$.

Определение 18. Через $|w|_a$ обозначается количество вхождений символа a в слово w.

Пример 19. Если $\Sigma = \{a, b, c\}$, то $|baaa|_a = 3$, $|baaa|_b = 1$ и $|baaa|_c = 0$.

Определение 20. Функция $f\colon K \to L$ называется *биекцией* (bijection), если каждый элемент множества L является образом ровно одного элемента множества K (относительно функции f).

Задачи

1.1.1. Сколько трёхбуквенных слов содержит язык $\{a, b, c\}^*$?

1.1.2. Сколько десятибуквенных слов содержит язык $\{a, b\}^*$?

1.1.3. Равны ли языки $\{a^m b^m \mid m \geqslant 0\}$ и $\{(ab)^m \mid m \geqslant 0\}$?

1.1.4. Равны ли языки $\{a^m b^m \mid m > 0\}$ и $\{a^k b^k \mid k \geqslant 1\}$?

1.1.5. Равны ли языки $\{a^{m+1}ba^m \mid m \geqslant 0\}$ и $\{a^k ba^m \mid k = m + 1\}$?

1.1.6. Верно ли, что $\{a^{2m+1} \mid m \geqslant 0\} \subseteq \{a^m \mid m \geqslant 0\}$?

1.1.7. Верно ли, что $\{a^{2m} \mid m \geqslant 0\} \subseteq \{a^k b^n a^k \mid k \geqslant 0,\ n \geqslant 0\}$?

1.1.8. Верно ли, что $\{a^m b^m \mid m > 0\} \subseteq \{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$?

1.1.9. Сколько пятибуквенных слов содержит язык $L = \{v_1 a v_2 c v_3 b v_4 \mid v_1, v_2, v_3, v_4 \in \{a, b\}^*,\ v_1 \neq v_4,\ v_2 \neq v_3\}$ над алфавитом $\{a, b, c\}^*$?

1.2. Операции над языками

[20, 1.2], [6, с. 14, 35], [2, 0.2.1, 0.2.3], [24, 2.1], [23, 1.1], [7, 1.1], [5, с. 3—4, 16], [3, 9.1], [11, 1.3.3], [31, 1.7], [13, 1.3.1], [14, с. 262], [10, 2.2], [1, 3.3], [18, с. 38], [4, с. 105—106], [29, 2.4], [17, 5.2]

Определение 1. Пусть $L_1, L_2 \subseteq \Sigma^*$. Тогда $L_1 \cdot L_2 \leftrightharpoons \{xy \mid x \in L_1,\ y \in L_2\}$. Язык $L_1 \cdot L_2$ называется *конкатенацией* языков L_1 и L_2.

Пример 2. Если $L_1 = \{a, abb\}$ и $L_2 = \{bbc, c\}$, то $L_1 \cdot L_2 = \{ac, abbc, abbbbc\}$.

Определение 3. Пусть $L \subseteq \Sigma^*$. Тогда

$$
\begin{aligned}
&L^0 \leftrightharpoons \{\varepsilon\},\\
&L^n \leftrightharpoons \underbrace{L \cdot \ldots \cdot L}_{n \text{ раз}}, \text{ если } n > 0.
\end{aligned}
$$

Пример 4. Если $L = \{a^k ba^l \mid 0 < k < l\}$, то $L^2 = \{a^k ba^l ba^m \mid 0 < k < l - 1,\ m > 1\}$.

Определение 5. *Итерацией* (Kleene closure) языка L (обозначение L^*) называется язык $\bigcup\limits_{n \in \mathbb{N}} L^n$. Эта операция называется также *звёздочкой Клини* (Kleene star, star operation).

Пример 6. Если $\Sigma = \{a, b\}$ и $L = \{aa, ab, ba, bb\}$, то $L^* = \{w \in \Sigma^* \mid |w| \text{ чётно}\}$.

Определение 7. *Обращением* или *зеркальным образом* (reversal) слова w (обозначается w^{R}) называется слово, в котором символы, составляющие слово w, идут в обратном порядке.

Пример 8. Если $w = baaca$, то $w^{\mathrm{R}} = acaab$.

Определение 9. Пусть $L \subseteq \Sigma^*$. Тогда $L^{\mathrm{R}} \leftrightharpoons \{w^{\mathrm{R}} \mid w \in L\}$. Язык L^{R} называется *обращением* языка L.

Определение 10. Говорят, что слово x — *префикс* (*начало*) слова y (обозначение $x \sqsubset y$), если $y = xu$ для некоторого слова u.

Пример 11. Очевидно, что $\varepsilon \sqsubset baa$, $b \sqsubset baa$, $ba \sqsubset baa$ и $baa \sqsubset baa$.

Определение 12. Говорят, что слово x — *суффикс* (*конец*) слова y (обозначение $x \sqsupset y$), если $y = ux$ для некоторого слова u.

Определение 13. Говорят, что слово x — *подслово* (substring) слова y, если $y = uxv$ для некоторых слов u и v.

Определение 14. Пусть $w \in \Sigma^*$. Тогда через $\mathrm{Subw}(w)$ обозначается множество, состоящее из всех подслов слова w.

Определение 15. Пусть $L \subseteq \Sigma^*$. Тогда через $\mathrm{Subw}(L)$ обозначается множество, состоящее из всех подслов слов языка L. Множество $\mathrm{Subw}(L)$ называется *множеством подслов* языка L.

Определение 16. Слово $a_1a_2 \ldots a_n$ (длины $n \geqslant 0$) называется *подпоследовательностью* (subsequence) слова y, если существуют такие слова u_0, $u_1, \ldots, u_n$, что $u_0a_1u_1a_2u_2 \ldots a_nu_n = y$.

Замечание 17. Все подслова слова y являются также подпоследовательностями слова y.

Задачи

1.2.1. Существуют ли такие языки L_1, L_2 и L_3, что языки $(L_1 \cup L_2) \cdot L_3$ и $(L_1 \cdot L_3) \cup (L_2 \cdot L_3)$ не совпадают?

1.2.2. Существуют ли такие языки L_1, L_2 и L_3, что языки $(L_1 \cap L_2) \cdot L_3$ и $(L_1 \cdot L_3) \cap (L_2 \cdot L_3)$ не совпадают?

1.2.3. Существует ли такой конечный язык L, что $L^* = L$?

1.2.4. Существует ли такой бесконечный язык L, что $L^* = L$?

1.2.5. Верно ли, что $\{v_1cv_2 \mid v_1 \in \{a,b\}^*, v_2 \in \{a,b\}^*, v_1 \neq v_2\} \subseteq$ $\subseteq \{w \in \{a,b,c\}^* \mid w^{\mathrm{R}} \neq w\}$?

1.2.6. Равны ли языки $\{a,b,c\}^* - \{v_1cv_1 \mid v_1 \in \{a,b\}^*\}$ и $\{v_1cv_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ v_1 \neq v_2\}$?

1.2.7. Равны ли языки

$$\{v_1cv_2 \mid v_1, v_2 \in \{a,b\}^*,\ |v_1| = |v_2|,\ v_1 \neq v_2\}$$

и

$$\{v_1av_2cv_3bv_4 \mid v_1,v_2,v_3,v_4 \in \{a,b\}^*,\ |v_1| = |v_3|,\ |v_2| = |v_4|\} \cup$$
$$\cup\, \{v_1bv_2cv_3av_4 \mid v_1,v_2,v_3,v_4 \in \{a,b\}^*,\ |v_1| = |v_3|,\ |v_2| = |v_4|\}?$$

1.2.8. Равны ли языки $\{v_1cv_2 \mid v_1,v_2 \in \{a,b\}^*,\ v_1^{\mathrm{R}} \neq v_2\}$ и

$$\{v_1av_2cv_3bv_4 \mid v_1,v_2,v_3,v_4 \in \{a,b\}^*,\ |v_1| = |v_4|\} \cup$$
$$\cup\, \{v_1bv_2cv_3av_4 \mid v_1,v_2,v_3,v_4 \in \{a,b\}^*,\ |v_1| = |v_4|\}?$$

1.2.9. Обозначим $L = \{aa, ab, ba, bb\}$. Равны ли языки L^* и $(\{a,b\} \cdot L^* \cdot \{a,b\}) \cup \{\varepsilon\}$?

1.2.10. Обозначим $L = \{aa, ab, bb\}$. Равны ли языки L^* и $(\{a,b\} \cdot L^* \cdot \{a,b\}) \cup \{\varepsilon\}$?

1.3. Гомоморфизмы

[20, 1.3], [23, с. 10], [6, с. 57], [2, 0.2.3], [27, 4.2.3, 4.2.4], [18, I.9.2], [7, 1.1], [14, с. 259], [31, с. 85]

Определение 1. Пусть Σ_1 и Σ_2 — алфавиты. Если отображение $h\colon \Sigma_1^* \to \Sigma_2^*$ удовлетворяет условию $h(x \cdot y) = h(x) \cdot h(y)$ для всех слов $x \in \Sigma_1^*$ и $y \in \Sigma_1^*$, то отображение h называется *гомоморфизмом* (*морфизмом*).

Замечание 2. Можно доказать, что если h — гомоморфизм, то $h(\varepsilon) = \varepsilon$.

Пример 3. Пусть $\Sigma_1 = \{a,b\}$ и $\Sigma_2 = \{c\}$. Тогда отображение $h\colon \Sigma_1^* \to \Sigma_2^*$, заданное равенством $h(w) = c^{2|w|}$, является гомоморфизмом.

Замечание 4. Каждый гомоморфизм однозначно определяется своими значениями на однобуквенных словах.

Определение 5. Если $h\colon \Sigma_1^* \to \Sigma_2^*$ — гомоморфизм и $L \subseteq \Sigma_1^*$, то через $h(L)$ обозначается язык $\{h(w) \mid w \in L\}$.

Пример 6. Пусть $\Sigma = \{a,b\}$ и гомоморфизм $h\colon \Sigma^* \to \Sigma^*$ задан равенствами $h(a) = abba$ и $h(b) = \varepsilon$. Тогда $h(\{baa, bb\}) = \{abbaabba, \varepsilon\}$.

Определение 7. Если $h\colon \Sigma_1^* \to \Sigma_2^*$ — гомоморфизм и $L \subseteq \Sigma_2^*$, то через $h^{-1}(L)$ обозначается язык $\{w \in \Sigma_1^* \mid h(w) \in L\}$.

Пример 8. Рассмотрим алфавит $\Sigma = \{a, b\}$. Пусть гомоморфизм $h\colon \Sigma^* \to \Sigma^*$ задан равенствами $h(a) = ab$ и $h(b) = abb$. Тогда $h^{-1}(\{\varepsilon, abbb, abbab, ababab\}) = \{\varepsilon, ba, aaa\}$.

Задачи

1.3.1. Пусть гомоморфизм $h\colon \{a, b\}^* \to \{c, d, e\}^*$ задан соотношениями $h(a) = edc$, $h(b) = \varepsilon$. Описать язык $h(L)$, где $L = \{w \in \{a, b\}^* \mid |w|_a \leqslant |w|_b\}$

1.3.2. Пусть гомоморфизм $h\colon \{a, b, c\}^* \to \{a, b, c\}^*$ задан соотношениями $h(a) = abc$, $h(b) = \varepsilon$, $h(c) = bb$. Описать язык $h(L)$, где $L = \{w \in \{a, b\}^* \mid |w|_a = 2\}$

1.3.3. Пусть гомоморфизм $h\colon \{c, d, f\}^* \to \{a, b\}^*$ задан соотношениями $h(c) = a$, $h(d) = b$, $h(f) = \varepsilon$. Описать язык $h^{-1}(L)$, где $L = \{w \in \{a, b\}^* \mid |w|_a = |w|_b\}$

1.3.4. Пусть гомоморфизм $h\colon \{c, d, f\}^* \to \{a, b\}^*$ задан соотношениями $h(c) = a$, $h(d) = ab$, $h(f) = aba$. Описать язык $h^{-1}(L)$, где $L = \{w \in \{a, b\}^* \mid |w|_a = |w|_b\}$

1.3.5. Пусть гомоморфизм $h\colon \{a, b, c\}^* \to \{a, b, c\}^*$ задан соотношениями $h(a) = b$, $h(b) = a$, $h(c) = ab$. Описать язык $h^{-1}(L)$, где $L = \{a^n b^n c^k \mid n \geqslant 0,\ k \geqslant 0\}$

1.3.6. Пусть гомоморфизм $h\colon \{a, b, c\}^* \to \{a, b, c\}^*$ задан соотношениями $h(a) = abbcc$, $h(b) = b$, $h(c) = a$. Описать язык $h^{-1}(L)$, где $L = \{a^n b^n c^k \mid n \geqslant 0,\ k \geqslant 2\}$

1.3.7. Пусть гомоморфизм $h\colon \{a, b, c\}^* \to \{a, b, c\}^*$ задан соотношениями $h(a) = caaabbbc$, $h(b) = a$, $h(c) = bb$. Описать язык $h^{-1}(L)$, где $L = \{a^n b^n c^k \mid n \geqslant 3,\ k \geqslant 1\}$

1.4. Порождающие грамматики

[20, 1.4], [6, 1.1], [23, 2.1], [2, 2.1.2], [18, I.2.2], [7, 1.2], [15, с. 159—161], [4, с. 32—36], [8, с. 34—48, 177], [9, с. 354—355, 367—370], [24, 2.3.1], [5, с. 4—5], [25, с. 12—13], [26, 1.12], [22, с. 28—30], [13, 1.3.2], [14, с. 264—268], [10, 2.10], [17, 6.3—6.5], [31, 4.6]

Определение 1. *Порождающей грамматикой* (*грамматикой типа* 0, generative grammar, rewrite grammar) называется четвёрка $G \leftrightharpoons \langle N, \Sigma, P, S\rangle$, где N и Σ — конечные алфавиты, $N \cap \Sigma = \varnothing$, $P \subset (N \cup \Sigma)^+ \times (N \cup \Sigma)^*$, P конечно и $S \in N$.

Здесь Σ — *основной алфавит* (*терминальный алфавит*), его элементы называются *терминальными символами* или *терминалами* (terminal), N — *вспомогательный алфавит* (*нетерминальный алфавит*), его элементы называются *нетерминальными символами*, *нетерминалами*, *вспомогательными символами* или *переменными* (nonterminal, variable), S — *начальный символ* (*аксиома*, start symbol). Пары $(\alpha, \beta) \in P$ называются *правилами подстановки*, просто *правилами* или *продукциями* (rewriting rule, production) и записываются в виде $\alpha \to \beta$.

Пример 2. Пусть $N = \{S\}$, $\Sigma = \{a, b, c\}$, $P = \{S \to acSbcS,\ cS \to \varepsilon\}$. Тогда $\langle N, \Sigma, P, S\rangle$ является порождающей грамматикой.

Замечание 3. Будем обозначать элементы множества Σ строчными буквами из начала латинского алфавита, а элементы множества N — заглавными латинскими буквами. Обычно в примерах мы будем задавать грамматику в виде списка правил, подразумевая, что алфавит N составляют все заглавные буквы, встречающиеся в правилах, а алфавит Σ — все строчные буквы, встречающиеся в правилах. При этом правила порождающей грамматики записывают в таком порядке, что левая часть первого правила есть начальный символ S.

Замечание 4. Для обозначения n правил с одинаковыми левыми частями $\alpha \to \beta_1, \ldots, \alpha \to \beta_n$ часто используют сокращённую запись $\alpha \to \beta_1 \mid \ldots \mid \beta_n$.

Определение 5. Пусть дана грамматика G. Пишем $\varphi \underset{G}{\Rightarrow} \psi$, если $\varphi = \eta\alpha\theta$, $\psi = \eta\beta\theta$ и $(\alpha \to \beta) \in P$ для некоторых слов α, β, η, θ в алфавите $N \cup \Sigma$. Если $\varphi \underset{G}{\Rightarrow} \psi$, то говорят, что слово ψ *непосредственно выводимо* из слова φ.

Замечание 6. Когда из контекста ясно, о какой грамматике идёт речь, вместо $\underset{G}{\Rightarrow}$ можно писать просто $\Rightarrow$.

Пример 7. Рассмотрим грамматику

$$G = \langle \{S\}, \{a, b, c\}, \{S \to acSbcS,\ cS \to \varepsilon\}, S\rangle.$$

Тогда $cSacS \underset{G}{\Rightarrow} cSa$.

Определение 8. Если $\omega_0 \underset{G}{\Rightarrow} \omega_1 \underset{G}{\Rightarrow} \ldots \underset{G}{\Rightarrow} \omega_n$, где $n \geqslant 0$, то пишем $\omega_0 \overset{*}{\underset{G}{\Rightarrow}} \omega_n$ и говорим, что слово ω_n *выводимо* из слова ω_0. Другими словами, бинарное отношение $\overset{*}{\underset{G}{\Rightarrow}}$ является рефлексивным, транзитивным замыканием бинарного отношения $\underset{G}{\Rightarrow}$, определённого на множестве $(N \cup \Sigma)^*$.

При этом последовательность слов $\omega_0, \omega_1, \ldots, \omega_n$ называется *выводом* (derivation) слова ω_n из слова ω_0 в грамматике G. Число n называется *длиной* (*количеством шагов*) этого вывода.

Пример 9. Пусть $G = \langle \{S\}, \{a, b\}, \{S \to aS,\ \ S \to b\}, S \rangle$. Тогда $aS \overset{*}{\underset{G}{\Rightarrow}} aaaaS$. Длина этого вывода — 3.

Определение 10. Язык, *порождаемый* грамматикой G, — это множество $L(G) \leftrightharpoons \{\omega \in \Sigma^* \mid S \overset{*}{\underset{G}{\Rightarrow}} \omega\}$. Будем также говорить, что грамматика G *порождает* (generates) язык $L(G)$.

Пример 11. Если $G = \langle \{S\}, \{a, b\}, \{S \to aS,\ \ S \to bb\}, S \rangle$, то $L(G) = \{a^n bb \mid n \geqslant 0\}$.

Определение 12. Две грамматики *эквивалентны*, если они порождают один и тот же язык.

Пример 13. Грамматика $S \to abS$, $S \to a$ и грамматика $T \to aU$, $U \to baU$, $U \to \varepsilon$ эквивалентны.

Задачи

1.4.1. Принадлежит ли слово $aaaaabbbabb$ языку, порождаемому грамматикой $F \to aFbbb$, $F \to aFba$, $F \to aFa$, $F \to aFabb$, $F \to a$?

1.4.2. Принадлежит ли слово $abababa$ языку, порождаемому грамматикой $J \to aa$, $J \to FbF$, $F \to aJ$, $F \to Jb$, $F \to bFa$, $F \to \varepsilon$?

1.4.3. Является ли бесконечным язык, порождаемый грамматикой $S \to VZ$, $ZV \to ZVZV$, $T \to aa$, $T \to bb$, $V \to aTb$, $V \to bTa$, $Z \to aab$, $Z \to b$,

1.4.4. Описать язык над алфавитом $\{a, b, c\}$, порождаемый грамматикой $S \to abSc$, $S \to abc$, $ab \to ba$, $ba \to ab$, $ac \to ca$, $bc \to cb$.

1.4.5. Описать язык над алфавитом $\{a, b, c\}$, порождаемый грамматикой $S \to abSc$, $S \to bSc$, $S \to Sc$, $S \to bcc$, $ab \to ba$, $ba \to ab$, $ac \to ca$, $bc \to cb$.

1.4.6. Описать язык над алфавитом $\{a, b, c\}$, порождаемый грамматикой $T \to TAa$, $T \to TBb$, $T \to U$, $aA \to Aa$, $aB \to Ba$, $bA \to Ab$, $bB \to Bb$, $UA \to aU$, $UB \to bU$, $U \to c$.

1.4.7. Описать язык над алфавитом $\{a, b\}$, порождаемый грамматикой $F \to aABF$, $F \to bb$, $Aa \to aA$, $Ba \to aB$, $Bb \to bA$, $Ab \to ba$.

1.4.8. Описать язык над алфавитом $\{a, b\}$, порождаемый грамматикой $S \to SCB$, $S \to T$, $BC \to CBaa$, $aB \to Ba$, $aC \to Ca$, $TC \to T$, $T \to R$, $RB \to R$, $R \to b$.

1.4.9. Описать язык над алфавитом $\{a, b\}$, порождаемый грамматикой $S \to SCB$, $S \to T$, $BC \to CBaa$, $aB \to Ba$, $aC \to Ca$, $TC \to T$, $T \to R$, $RB \to Raaa$, $R \to b$.

1.4.10. Описать язык над алфавитом $\{a, b\}$, порождаемый грамматикой $S \to bbSCB$, $S \to SCB$, $S \to bbT$, $S \to bT$, $BC \to CBaa$, $aB \to Ba$, $aC \to Ca$, $TC \to Ta$, $T \to R$, $RB \to Raa$, $R \to a$.

1.4.11. Описать язык над алфавитом $\{a, b\}$, порождаемый грамматикой $S \to aaSCB$, $S \to SCB$, $S \to aaT$, $S \to aT$, $BC \to CBaa$, $aB \to Ba$, $aC \to Ca$, $T \to J$, $JJ \to bb$, $TC \to Ta$, $T \to R$, $RB \to Raa$, $R \to a$.

1.4.12. Эквивалентны ли грамматика $S \to TT$, $T \to bTa$, $T \to a$ и грамматика $S \to RT$, $R \to bRa$, $R \to T$, $T \to bTa$, $T \to a$?

1.4.13. Эквивалентны ли грамматика $S \to SabaS$, $S \to b$ и грамматика $S \to baSaS$, $S \to b$?

1.5. Классы грамматик

[20, 1.5], [6, с. 23—24, 78—79], [2, 2.1.3, с. 191], [23, 2.1, с. 94], [18, I.2.3, I.4.7], [7, 1.2, 1.3], [4, с. 39—45], [9, с. 361—367], [24, 2.3.2], [5, с. 6—8], [26, 1.12], [14, с. 268—271], [8, с. 54, 63, 178], [13, 1.3.2], [19, 5.3, 5.4, 10.2], [10, 2.10], [16, 5.2.1]

Определение 1. *Контекстной грамматикой* (*контекстно-зависимой грамматикой, грамматикой непосредственно составляющих, НС-грамматикой, грамматикой типа* 1, context-sensitive grammar, phrase-structure grammar) называется порождающая грамматика, каждое правило которой имеет вид $\eta A\theta \to \eta\alpha\theta$, где $A \in N$, $\eta \in (N \cup \Sigma)^*$, $\theta \in (N \cup \Sigma)^*$, $\alpha \in (N \cup \Sigma)^+$.

Пример 2. Грамматика $S \to TS$, $S \to US$, $S \to b$, $Tb \to Ab$, $A \to a$, $TA \to AAT$, $UAb \to b$, $UAAA \to AAU$ не является контекстной (последние три правила не имеют требуемого вида).

Определение 3. *Контекстно-свободной грамматикой* (*бесконтекстной грамматикой, КС-грамматикой, грамматикой типа* 2, context-free grammar) называется порождающая грамматика, каждое правило которой имеет вид $A \to \alpha$, где $A \in N$, $\alpha \in (N \cup \Sigma)^*$.

Пример 4. Грамматика $S \to ASTA$, $S \to AbA$, $A \to a$, $bT \to bb$, $AT \to UT$, $UT \to UV$, $UV \to TV$, $TV \to TA$ является контекстной, но не контекстно-свободной (последние пять правил не имеют требуемого вида).

Определение 5. *Линейной грамматикой* (linear grammar) называется порождающая грамматика, каждое правило которой имеет вид $A \to u$ или $A \to uBv$, где $A \in N$, $u \in \Sigma^*$, $v \in \Sigma^*$, $B \in N$.

Пример 6. Грамматика $S \to TT$, $T \to cTT$, $T \to bT$, $T \to a$ является контекстно-свободной, но не линейной (первые два правила не имеют требуемого вида).

Определение 7. *Праволинейной грамматикой* (*рациональной грамматикой, грамматикой типа* 3, right-linear grammar) называется порождающая грамматика, каждое правило которой имеет вид $A \to u$ или $A \to uB$, где $A \in N$, $u \in \Sigma^*$, $B \in N$.

Пример 8. Грамматика $S \to aSa$, $S \to T$, $T \to bT$, $T \to \varepsilon$ является линейной, но не праволинейной (первое правило не имеет требуемого вида).

Пример 9. Грамматика $S \to T$, $U \to abba$ праволинейная. Она порождает язык $\varnothing$.

Пример 10. Рассмотрим однобуквенный алфавит $\Sigma = \{a\}$. При любых фиксированных $k \in \mathbb{N}$ и $m \in \mathbb{N}$ язык $\{a^{k+mn} \mid n \in \mathbb{N}\}$ является праволинейным. Он задаётся грамматикой $S \to a^m S$, $S \to a^k$.

Определение 11. Правила вида $\alpha \to \varepsilon$ называются *ε-правилами* или *эпсилон-правилами*.

Лемма 12. *Каждая праволинейная грамматика является линейной. Каждая линейная грамматика является контекстно-свободной. Каждая контекстно-свободная грамматика без*

ε-правил является контекстной грамматикой.

Определение 13. Классы грамматик типа 0, 1, 2 и 3 образуют *иерархию Хомского* (Chomsky hierarchy).

Определение 14. Язык называется *языком типа* 0 (*контекстным языком*, *контекстно-свободным языком*, *линейным языком*, *праволинейным языком*), если он порождается некоторой грамматикой типа 0 (соответственно контекстной грамматикой, контекстно-свободной грамматикой, линейной грамматикой, праволинейной грамматикой). Контекстно-свободные языки называются также *алгебраическими языками*.

Пример 15. Пусть $\Sigma = \{a, b, c\}$. Язык $\{ucu^{\mathrm{R}} \mid u \in \{a, b\}^*\}$ является линейным, так как он порождается грамматикой $S \to aSa$, $S \to bSb$, $S \to c$.

Пример 16. Пусть дан произвольный алфавит

$$\Sigma = \{a_1, \ldots, a_n\}.$$

Тогда язык Σ^* является праволинейным, так как он порождается грамматикой $S \to \varepsilon$, $S \to a_1S, \ldots,$ $S \to a_nS$.

Задачи

1.5.1. Какому классу принадлежит грамматика $S \to aS$, $S \to bS$, $S \to \varepsilon$?

1.5.2. Какому классу принадлежит грамматика $S \to abSba$, $S \to a$?

1.5.3. Какому классу принадлежит грамматика $S \to aSbSa$, $S \to a$?

1.5.4. Какому классу принадлежит грамматика $K \to KT$, $T \to Ba$, $KB \to B$, $B \to b$?

1.5.5. Какому классу принадлежит грамматика $S \to aSa$, $S \to T$, $T \to Tb$, $T \to \varepsilon$?

1.5.6. Какому классу принадлежит грамматика $F \to aFb$, $F \to bFa$, $F \to FF$, $F \to c$?

1.5.7. Какому классу принадлежит грамматика $S \to TaV$, $T \to bV$, $V \to cVc$, $V \to \varepsilon$, $ac \to c$?

В задачах 1.5.8—1.5.17 надо найти линейную грамматику, порождающую данный язык.

1.5.8. $\{a^nba^n \mid n \geqslant 1\}$.

1.5.9. $\{a^m b^n \mid m = n\}$.

1.5.10. $\{a^m b^n \mid m = 2n\}$.

1.5.11. $\{a^m b^n \mid 3m = 2n\}$.

1.5.12. $\{a^m b^n c^n d^m \mid m \geqslant 0,\ n \geqslant 1\}$.

1.5.13. $\{a^n b^{2m} a b^m a^{3n} \mid m \geqslant 0,\ n \geqslant 0\}$.

1.5.14. $\{a^m b^n \mid 0 \leqslant n \leqslant m\}$.

1.5.15. $\{a^m b^n \mid 0 \leqslant n < m\}$.

1.5.16. $\{a^m b^n \mid 1 \leqslant m < n\}$.

1.5.17. $\{a^m b^n \mid 0 \leqslant n \leqslant m \leqslant 2n\}$.

1.5.18. Найти контекстно-свободную грамматику, порождающую язык $\{a^m b^m c^n d^n \mid m \geqslant 0,\ n \geqslant 1\}$.

1.5.19. Существует ли линейная грамматика с пятью правилами, порождающая конечный язык над алфавитом $\{a, b\}$, содержащий ровно 6 слов?

1.5.20. Существует ли контекстно-свободная грамматика с пятью правилами, порождающая конечный язык над алфавитом $\{a, b\}$, содержащий ровно 98 слов?

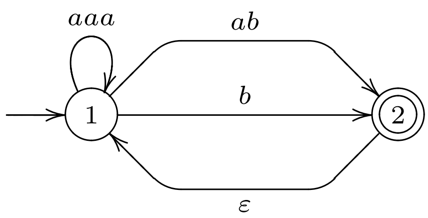
1.6. Автоматные языки

[20, 2.1, 2.3—2.4, 2.6, 3.1—3.2], [6, 2.1—2.2], [23, 2.1—2.3, с. 58], [2, 2.2.3—2.2.4, 2.3.3], [27, 2.2, 2.3.1—2.3.4, 2.5, 3.2.3, 4.2.1—4.2.2, 4.2.5], [18, I.3.1, I.3.3—I.3.5], [7, 5.1—5.2], [9, с. 391—393, 408—413], [4, с. 104—108], [11, 10.2, 10.3.1, 10.4.1—10.4.3], [26, 0.1—0.2, 1.1, 1.5—1.6], [22, с. 40—53]

Определение 1. *Конечный автомат* (finite automaton, finite-state machine) — это пятёрка $M = \langle Q, \Sigma, \Delta, I, F\rangle$, где Σ — конечный *входной алфавит* (или просто *алфавит*) данного конечного автомата, Q и Δ — конечные множества, $\Delta \subseteq Q \times \Sigma^* \times Q$, $I \subseteq Q$, $F \subseteq Q$. Элементы Q называются *состояниями* (state), элементы I — *начальными* (initial) состояниями, элементы F — *заключительными* или *допускающими* (final, accepting) состояниями. Если $\langle p, x, q\rangle \in \Delta$, то $\langle p, x, q\rangle$ называется *переходом* (transition) из p в q, а слово x — *меткой* (label) этого перехода.

Замечание 2. Конечные автоматы можно изображать в виде *диаграмм состояний* (state diagram) (иногда их называют *диаграммами переходов* (transition diagram)). На диаграмме каждое состояние обозначается кружком, а переход — стрелкой. Стрелка из p в q, помеченная словом x, показывает, что $\langle p, x, q\rangle$ является переходом данного конечного автомата. Каждое начальное состояние распознаётся по ведущей в него короткой стрелке. Каждое допускающее состояние отмечается на диаграмме двойным кружком.

Пример 3. На диаграмме изображён конечный автомат $\langle Q, \Sigma, \Delta, I, F\rangle$, где $Q = \{1, 2\}$, $\Sigma = \{a, b\}$, $I = \{1\}$, $F = \{2\}$, $\Delta = \{\langle 1, aaa, 1\rangle,\ \langle 1, ab, 2\rangle,\ \langle 1, b, 2\rangle,\ \langle 2, \varepsilon, 1\rangle\}$.



Определение 4. *Путь* (path) конечного автомата — это кортеж $\langle q_0, e_1, q_1, e_2, \ldots, q_n\rangle$, где $n \geqslant 0$ и $e_i = \langle q_{i-1}, w_i, q_i\rangle \in \Delta$ для каждого i. При этом q_0 — *начало пути*, q_n — *конец пути*, n — *длина пути*, $w_1 \ldots w_n$ — *метка пути*.

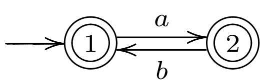
Определение 5. Путь называется *успешным*, если его начало принадлежит I, а конец принадлежит F.

Пример 6. Рассмотрим конечный автомат из примера 3. Путь $\langle 1,\ \langle 1,b,2\rangle,\ 2,\ \langle 2,\varepsilon,1\rangle,\ 1,\ \langle 1,aaa,1\rangle,\ 1,\ \langle 1,b,2\rangle,\ 2\rangle$ является успешным. Его метка — $baaab$. Длина этого пути — 4.

Определение 7. Конечный автомат M *допускает* (accepts, recognizes) слово w, если оно является меткой некоторого успешного пути.

Определение 8. Язык, *распознаваемый* конечным автоматом M, — это язык $L(M)$, состоящий из меток всех успешных путей (то есть из всех допускаемых данным автоматом слов). Будем также говорить, что автомат M *распознаёт* (recognizes, accepts) язык $L(M)$.

Пример 9. Пусть $M = \langle Q, \Sigma, \Delta, I, F\rangle$, где $Q = \{1, 2\}$, $\Sigma = \{a, b\}$, $I = \{1\}$, $F = \{1, 2\}$, $\Delta = \{\langle 1, a, 2\rangle,\ \langle 2, b, 1\rangle\}$.



Тогда $L(M) = \{(ab)^n \mid n \geqslant 0\} \cup \{(ab)^n a \mid n \geqslant 0\}$.

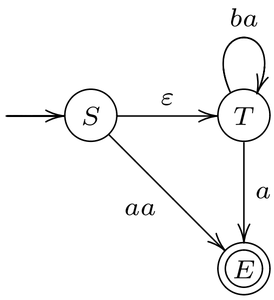
Определение 10. Два конечных автомата *эквивалентны*, если они распознают один и тот же язык.

Определение 11. Язык L называется *автоматным* (finite-state language), если существует конечный автомат, распознающий этот язык.

Теорема 12. *Язык является автоматным тогда и только тогда, когда он является праволинейным.*

Пример 13. Язык, распознаваемый конечным автоматом из примера 3, порождается грамматикой $K_1 \rightarrow aaaK_1$, $K_1 \rightarrow abK_2$, $K_1 \rightarrow bK_2$, $K_2 \rightarrow K_1$, $K_2 \rightarrow \varepsilon$.

Пример 14. Пусть $\Sigma = \{a, b\}$. Рассмотрим грамматику $S \rightarrow aa$, $S \rightarrow T$, $T \rightarrow baT$, $T \rightarrow a$. Она эквивалентна грамматике $S \rightarrow aaE, S \rightarrow T, T \rightarrow baT, T \rightarrow aE, E \rightarrow \varepsilon$. Язык, порождаемый этими грамматиками, распознаётся конечным автоматом $\langle Q, \Sigma, \Delta, I, F\rangle$, где $Q = \{S, T, E\}$, $I = \{S\}$, $F = \{E\}$ и $\Delta = \{\langle S, aa, E\rangle,\ \langle S, \varepsilon, T\rangle,\ \langle T, ba, T\rangle,\ \langle T, a, E\rangle\}$.



Определение 15. Конечный автомат $\langle Q, \Sigma, \Delta, I, F\rangle$ называется *детерминированным* (deterministic), если

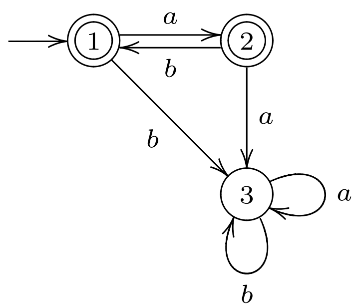
1) множество I содержит ровно один элемент;

2) для каждого перехода $\langle p, x, q\rangle \in \Delta$ выполняется равенство $|x| = 1$;

3) для любого символа $a \in \Sigma$ и для любого состояния $p \in Q$ существует не более одного состояния $q \in Q$ со свойством $\langle p, a, q\rangle \in \Delta$.

Определение 16. Детерминированный конечный автомат $M = \langle Q, \Sigma, \Delta, I, F\rangle$ называется *полным* (complete), если для каждого состояния $p \in Q$ и для каждого символа $a \in \Sigma$ найдётся такое состояние $q \in Q$, что $\langle p, a, q\rangle \in \Delta$.

Пример 17. Детерминированный конечный автомат из примера 9 эквивалентен следующему полному детерминированному конечному автомату.



Теорема 18. *Каждый автоматный язык распознаётся некоторым полным детерминированным конечным автоматом.*

Доказательство. Без ограничения общности можно предположить, что исходный язык задан конечным автоматом $\langle Q, \Sigma, \Delta, I, F\rangle$, содержащим только переходы с метками длины единица. Для любых $a \in \Sigma$ и $H \subseteq Q$ обозначим

$$\Delta_a(H) \risingdotseq \{q \in Q \mid \langle p, a, q\rangle \in \Delta \text{ для некоторого } p \in H\}.$$

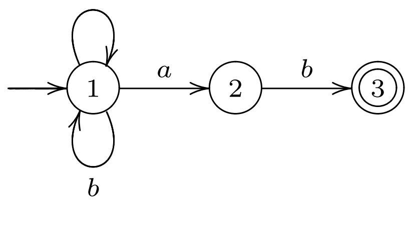
Обозначим через $\mathcal{P}(Q)$ множество всех подмножеств множества Q. Построим искомый полный детерминированный конечный автомат $\langle Q', \Sigma, \Delta', I', F'\rangle$, положив $Q' = \mathcal{P}(Q)$,

$$\Delta' = \{\langle H, a, \Delta_a(H)\rangle \mid H \subseteq Q,\ a \in \Sigma\},$$

$I' = \{I\}$ и $F' = \{H \subseteq Q \mid H \cap F \neq \varnothing\}$. □

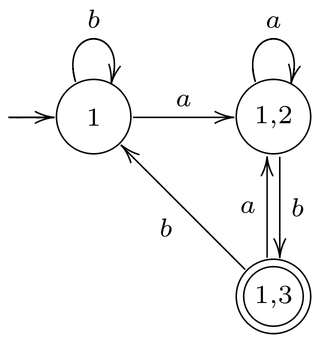
Пример 19. Пусть $\Sigma = \{a, b\}$. Рассмотрим конечный автомат $M = \langle\{1, 2, 3\}, \Sigma, \Delta, \{1\}, \{3\}\rangle$, где

$$\Delta = \{\langle 1, a, 1\rangle,\ \langle 1, b, 1\rangle,\ \langle 1, a, 2\rangle,\ \langle 2, b, 3\rangle\}.$$



Если применить конструкцию из доказательства теоремы 18 и затем удалить состояния, не достижимые из начального состояния, то получится полный детерминированный конечный автомат $M' = \langle \{\{1\},\{1,2\},\{1,3\}\}, \Sigma, \Delta', \{\{1\}\}, \{\{1,3\}\} \rangle$, где

$$\Delta' = \{\langle \{1\}, a, \{1,2\}\rangle,\ \langle \{1\}, b, \{1\}\rangle,\ \langle \{1,2\}, a, \{1,2\}\rangle,\\ \langle \{1,2\}, b, \{1,3\}\rangle,\ \langle \{1,3\}, a, \{1,2\}\rangle,\ \langle \{1,3\}, b, \{1\}\rangle\}.$$

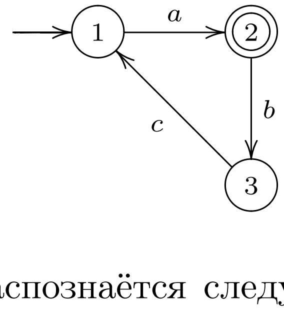


Теорема 20. *Класс автоматных языков замкнут относительно итерации, конкатенации, объединения, пересечения, дополнения и обращения.*

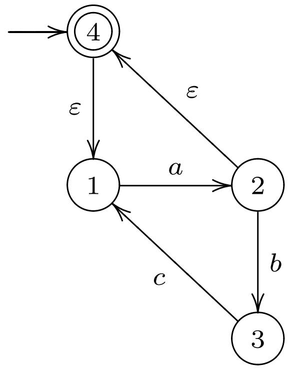
Доказательство. Для доказательства замкнутости относительно дополнения заметим, что если язык L распознаётся полным детерминированным конечным автоматом $\langle Q, \Sigma, \Delta, I, F\rangle$, то язык $\Sigma^* - L$ распознаётся конечным автоматом $\langle Q, \Sigma, \Delta, I, Q - F\rangle$.

Пересечение выражается через объединение и дополнение (закон де Моргана). □

Пример 21. Пусть $\Sigma = \{a, b, c\}$. Рассмотрим конечный автомат



Тогда язык $L(M_1)^*$ распознаётся следующим конечным автоматом.



Задачи

1.6.1. Является ли автоматным язык, порождаемый грамматикой $S \to abSabSab$, $S \to abab$?

1.6.2. Является ли автоматным язык, порождаемый грамматикой $S \to TcV$, $T \to bV$, $V \to aVa$, $V \to \varepsilon$, $ac \to a$?

1.6.3. Является ли автоматным язык, порождаемый грамматикой $S \to SaSaS$, $S \to b$?

1.6.4. Является ли язык $\{a^k b^m c^n \mid k \geqslant 0,\ m \geqslant 0,\ n \geqslant 0\}$ автоматным?

1.6.5. Найти конечный автомат для дополнения языка $\{a^k b^m c^n \mid k \geqslant 1,\ m \geqslant 1,\ n \geqslant 1\}$ над алфавитом $\{a, b, c\}$.

1.6.6. Найти конечный автомат для дополнения языка над алфавитом $\{a, b\}$, порождаемого грамматикой $S \to aS$, $S \to bbT$, $T \to aT$, $T \to bZ$, $Z \to aZ$, $Z \to \varepsilon$.

1.6.7. Найти конечный автомат для дополнения языка над алфавитом $\{a, b\}$, порождаемого грамматикой $K_1 \to aaaK_1$, $K_1 \to abK_2$, $K_1 \to bK_2$, $K_2 \to K_1$, $K_2 \to \varepsilon$.

1.6.8. Найти конечный автомат для дополнения языка над алфавитом $\{a, b\}$, порождаемого грамматикой $S \to aT$, $S \to bZ$, $T \to aZ$, $T \to bS$, $Z \to aS$, $Z \to bT$, $Z \to abT$, $S \to \varepsilon$.

1.6.9. Найти праволинейную грамматику для обращения языка, порождаемого грамматикой $S \to baS$, $S \to abcT$, $T \to aT$, $T \to bZ$, $Z \to aZ$, $Z \to acS$, $Z \to a$.

2. Неоднозначность в контекстно-свободных грамматиках

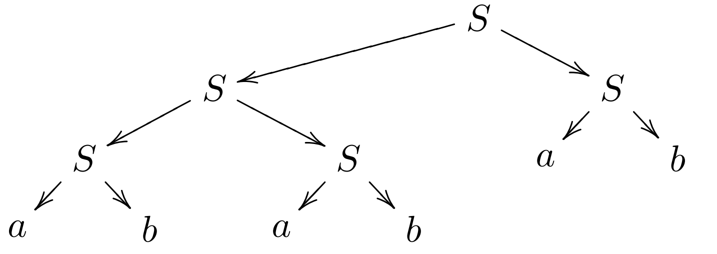
2.1. Деревья вывода

[20, 7.1], [6, 1.5], [27, 5.2.1, 5.2.2], [18, I.2.6], [25, с. 13—14, 31], [4, с. 45—56], [15, с. 294], [2, 0.5.1—0.5.4, 2.4.1], [9, с. 370—372], [24, 4.1], [5, с. 9], [11, 8.2.1], [7, 3.1], [31, 3.2], [22, с. 31—32], [1, 2.2, 4.2], [13, 3.1.1], [19, 7.1], [10, 2.4], [17, 6.6], [14, с. 277]

Определение 1. Выводам в контекстно-свободной грамматике соответствуют так называемые *деревья вывода* (*деревья разбора*, derivation tree, parse tree) — некоторые упорядоченные деревья, вершины которых помечены символами алфавита $N \cup \Sigma$. Корень дерева отвечает начальному символу. Каждому символу слова w_1, на которое заменяется начальный символ на первом шаге вывода, ставится в соответствие вершина дерева, и к ней проводится дуга из корня. Полученные таким образом непосредственные потомки корня упорядочены согласно порядку их меток в слове w_1. Для тех из полученных вершин, которые помечены символами из множества N, делается аналогичное построение, и т. д.

Кроной (yield) дерева вывода называется слово, записанное в вершинах, помеченных символами из алфавита Σ.

Пример 2. Рассмотрим контекстно-свободную грамматику $S \to SS$, $S \to ab$, $S \to aSb$. Выводу $S \Rightarrow SS \Rightarrow Sab \Rightarrow SSab \Rightarrow$ $\Rightarrow abSab \Rightarrow ababab$ соответствует следующее дерево вывода.



Крона этого дерева вывода — $ababab$.

Задачи

2.1.1. Перечислить все деревья вывода в грамматике $S \to a$, $S \to bR$, $R \to cTT$, $T \to a$, $T \to ba$.

2.1.2. Перечислить все деревья вывода в грамматике $S \to TV$, $S \to VT$, $T \to aVa$, $T \to \varepsilon$, $V \to bb$, $V \to \varepsilon$.

2.1.3. Перечислить все деревья вывода в грамматике $S \to aaT$, $S \to abT$, $T \to bbZ$, $Z \to a$, $Z \to b$, $Z \to \varepsilon$.

2.1.4. Перечислить все деревья вывода в грамматике $S \to VZ$, $T \to aa$, $T \to bb$, $V \to aTb$, $V \to bTa$, $Z \to aab$, $Z \to b$.

2.1.5. У скольких деревьев вывода в грамматике $S \to aSbSa$, $S \to a$ крона равна $aaabaabaaabaabaaa$?

2.1.6. Существует ли праволинейная грамматика без ε-правил, в которой некоторое слово имеет бесконечно много выводов?

2.1.7. Рассмотрим контекстно-свободную грамматику, где у каждого правила длина правой части равна 0 или 2. Пусть в некотором дереве вывода самый длинный ориентированный путь содержит 4 вершин. Может ли длина кроны этого дерева равняться 7?

2.1.8. Рассмотрим контекстно-свободную грамматику, где у каждого правила длина правой части равна 0 или 2. Пусть в некотором дереве вывода самый длинный ориентированный путь содержит 5 вершин. Может ли длина кроны этого дерева равняться 20?

2.2. Однозначные контекстно-свободные грамматики

[20, 7.2], [6, 1.6], [2, 2.4.1, 2.6.5], [27, 5.1.4, 5.2.3—5.2.6, 5.4], [18, I.4.1, I.4.7], [4, с. 56—58], [15, с. 307], [9, с. 372—374], [24, 4.1], [5, с. 8—10], [25, с. 14], [11, 8.2.3], [7, 3.1, 4.3], [31, 3.2], [32, 2.1], [22, с. 32], [1, 2.2, 4.2], [13, 3.1.1], [19, 7.1, 7.2], [10, 2.4], [17, 6.6], [14, с. 272—275]

Определение 1. Вывод в контекстно-свободной грамматике называется *левосторонним* (*левым*, leftmost derivation), если на каждом шаге вывода заменяется самое левое из всех вхождений вспомогательных символов (то есть каждый шаг вывода имеет вид $uA\theta \Rightarrow u\beta\theta$, где $(A \to \beta) \in P$, $u \in \Sigma^*$ и $\theta \in (N \cup \Sigma)^*$). Иногда в левосторонних выводах вместо $\Rightarrow$ пишут $\underset{\mathrm{lm}}{\Rightarrow}$. *Правосторонний* (*правый*) вывод определяется аналогично. В правосторонних выводах вместо $\Rightarrow$ пишут $\underset{\mathrm{rm}}{\Rightarrow}$.

Пример 2. Вывод

$$S \Rightarrow SS \Rightarrow Sab \Rightarrow SSab \Rightarrow abSab \Rightarrow ababab$$

из примера 2 на с. 24 не является левосторонним.

Лемма 3. *Для каждого слова, выводимого в контекстно-свободной грамматике, существует левосторонний вывод.*

Лемма 4. *Пусть G — контекстно-свободная грамматика над алфавитом Σ. Пусть $w \in \Sigma^*$. Тогда существует взаимно-однозначное соответствие между левосторонними выводами слова w в грамматике G и деревьями вывода в грамматике G, кроной которых является w.*

Пример 5. Рассмотрим дерево вывода из примера 2 на с. 24. Ему соответствует левосторонний вывод

$$S \underset{\mathrm{lm}}{\Rightarrow} SS \underset{\mathrm{lm}}{\Rightarrow} SSS \underset{\mathrm{lm}}{\Rightarrow} abSS \underset{\mathrm{lm}}{\Rightarrow} ababS \underset{\mathrm{lm}}{\Rightarrow} ababab.$$

Определение 6. Контекстно-свободная грамматика называется *неоднозначной* (ambiguous), если существует слово, которое имеет два или более левосторонних вывода (устаревший термин — *неопределённая* грамматика). В противном случае контекстно-свободная грамматика называется *однозначной* (unambiguous).

Пример 7. Контекстно-свободная грамматика из примера 2 на с. 24 неоднозначна. Слово $ababab$ имеет два левосторонних вывода:

$$S \underset{\mathrm{lm}}{\Rightarrow} SS \underset{\mathrm{lm}}{\Rightarrow} SSS \underset{\mathrm{lm}}{\Rightarrow} abSS \underset{\mathrm{lm}}{\Rightarrow} ababS \underset{\mathrm{lm}}{\Rightarrow} ababab$$

и

$$S \underset{\mathrm{lm}}{\Rightarrow} SS \underset{\mathrm{lm}}{\Rightarrow} abS \underset{\mathrm{lm}}{\Rightarrow} abSS \underset{\mathrm{lm}}{\Rightarrow} ababS \underset{\mathrm{lm}}{\Rightarrow} ababab.$$

Пример 8. Пусть $\Sigma = \{\boldsymbol{p}, \#, \mathbf{)}, \mathbf{(}, \neg, \wedge, \vee\}$. Контекстно-свободная грамматика $S \to S{\vee}D$, $S \to D$, $D \to D{\wedge}C$, $D \to C$,

$C \to \neg C$, $C \to (S)$, $C \to V$, $V \to V\#$, $V \to \boldsymbol{p}$ порождает множество всех булевых формул, составленных из переменных $\boldsymbol{p}$, $\boldsymbol{p}\#$, $\boldsymbol{p}\#\#$, $\ldots$ с помощью скобок и операторов $\neg$, $\wedge$ и $\vee$. Эта грамматика является однозначной.

Определение 9. Контекстно-свободный язык называется *существенно неоднозначным* (inherently ambiguous), если каждая контекстно-свободная грамматика, порождающая этот язык, является неоднозначной.

Пример 10. Язык, порождаемый контекстно-свободной грамматикой из примера 2 на с. 24, не является существенно неоднозначным. Он порождается однозначной грамматикой $S \to TS$, $S \to T$, $T \to ab$, $T \to aSb$.

Пример 11. Пусть $\Sigma = \{a, b, c\}$. Контекстно-свободный язык $\{a^k b^m c^n \mid k = m \text{ или } m = n\}$ является существенно неоднозначным. Доказательство этого факта можно найти в [2, с. 234—236].

Задачи

2.2.1. Однозначна ли контекстно-свободная грамматика $K \to a$, $K \to b$, $K \to K{+}K$, $K \to K{-}K$ с алфавитом $\Sigma = \{a, b, +, -\}$?

2.2.2. Однозначна ли контекстно-свободная грамматика $K \to a$, $K \to K + K$ с алфавитом $\Sigma = \{a, +\}$?

2.2.3. Однозначна ли контекстно-свободная грамматика $K \to a$, $K \to b$, $K \to +KK$, $K \to -KK$ с алфавитом $\Sigma = \{a, b, +, -\}$?

2.2.4. Однозначна ли контекстно-свободная грамматика $K \to a$, $K \to b$, $K \to (K{+}K)$, $K \to (K{-}K)$ с алфавитом $\Sigma = \{a, b, +, -,), (\}$?

2.2.5. Однозначна ли контекстно-свободная грамматика $K \to K{+}F$, $K \to K{-}F$, $K \to F$, $F \to a$, $F \to b$, $F \to (K)$ с алфавитом $\Sigma = \{a, b, +, -,), (\}$?

2.2.6. Однозначна ли контекстно-свободная грамматика $S \to aSbS$, $S \to bSaS$, $S \to \varepsilon$?

2.2.7. Однозначна ли контекстно-свободная грамматика $E \to EcE$, $E \to EdE$, $E \to aEb$, $E \to i$?

2.2.8. Однозначна ли контекстно-свободная грамматика $E \to EcT$, $E \to T$, $T \to TdF$, $T \to F$, $F \to aEb$, $F \to i$?

2.2.9. Однозначна ли контекстно-свободная грамматика $S \to T$, $S \to V$, $T \to aabTa$, $T \to aTaa$, $T \to daaTdb$, $T \to c$, $V \to aVa$, $V \to bVb$, $V \to dVd$, $V \to aca$, $V \to bcb$, $V \to dcd$?

2.2.10. Однозначна ли контекстно-свободная грамматика $S \to TV$, $S \to VT$, $T \to aVa$, $T \to bVb$, $V \to aTa$, $V \to bTb$, $T \to \varepsilon$?

2.2.11. Однозначна ли контекстно-свободная грамматика $S \to TV$, $S \to VT$, $T \to aVa$, $T \to bVb$, $V \to aTa$, $V \to bTb$, $T \to c$?

2.2.12. Найти однозначную контекстно-свободную грамматику, эквивалентную грамматике $S \to aS$, $S \to aSb$, $S \to c$.

2.2.13. Найти однозначную контекстно-свободную грамматику, эквивалентную грамматике $S \to aSbb$, $S \to aSb$, $S \to c$.

2.2.14. Найти однозначную контекстно-свободную грамматику, эквивалентную грамматике $S \to aSbbbb$, $S \to aaaSbb$, $S \to c$.

2.2.15. Найти однозначную контекстно-свободную грамматику, эквивалентную грамматике $S \to aSbb$, $S \to aSb$, $S \to aaSb$, $S \to c$.

2.2.16. Найти однозначную контекстно-свободную грамматику, эквивалентную грамматике $S \to aSaaaa$, $S \to aSaa$, $S \to aaSa$, $S \to b$.

2.2.17. Является ли существенно неоднозначным язык $\{a^k b^m \mid 1 \leqslant k \leqslant 2m\}$?

2.2.18. Является ли существенно неоднозначным язык $\{a^k b^l a^m b^n \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0,\ n \geqslant 0\}$?

2.2.19. Найти осмысленную однозначную грамматику для арифметических выражений с алфавитом $\{a, b, +, \times, (,)\}$

2.3. Языки Дика и Лукасевича

[20, 7.4], [23, с. 103, 116], [15, с. 293—295], [16, 5.1.3], [2, с. 239], [7, 6.5], [11, 15.1]

Определение 1. *Языком Дика* (Dyck language) над $2n$ буквами называется контекстно-свободный язык над алфавитом $\{a_1, b_1, a_2, b_2, \ldots, a_n, b_n\}$, порождаемый грамматикой $S \to \varepsilon$, $S \to a_1 S b_1 S, \ldots$, $S \to a_n S b_n S$.

Замечание 2. Словами этого языка являются последовательности правильно вложенных скобок n типов (если считать символ a_i левой скобкой, а символ b_i соответствующей правой скобкой).

Теорема 3. *Слово w над алфавитом $\{a, b\}$ выводится в грамматике $S \to \varepsilon$, $S \to aSbS$ тогда и только тогда, когда*

$$|w|_a = |w|_b$$

и для всех слов $x \sqsubset w$ выполняется неравенство

$$|x|_a \geqslant |x|_b.$$

Теорема 4. *При любом положительном целом n грамматика из определения 1 является однозначной.*

Определение 5. *Языком Лукасевича* (Lukasiewicz language) над $n+1$ буквами называется контекстно-свободный язык над алфавитом $\{a_0, a_1, \ldots, a_n\}$, порождаемый грамматикой

$S \to a_0$, $S \to a_1S$, $S \to a_2SS$,$\ldots$, $S \to a_nS^n$.

Теорема 6. *Слово w над алфавитом $\{a_0, a_1 \ldots, a_n\}$ принадлежит языку Лукасевича над $n+1$ буквами тогда и только тогда, когда*

$$\sum_{i=0}^{n} (i-1)|w|_{a_i} = -1$$

и для всех слов $x \sqsubset w$, кроме $x = w$, выполняется неравенство

$$\sum_{i=0}^{n} (i-1)|x|_{a_i} \geqslant 0.$$

Теорема 7. *При любом $n \in \mathbb{N}$ грамматика из определения 5 является однозначной.*

Задачи

2.3.1. Принадлежит ли слово $aababb$ языку, порождаемому грамматикой $S \to aSbS$, $S \to \varepsilon$?

2.3.2. Принадлежит ли слово $bccbaacaba$ языку, порождаемому грамматикой $S \to a$, $S \to bS$, $S \to cSS$?

2.3.3. Принадлежит ли слово $cacbcbbacacccaaaacabcaa$ языку, порождаемому грамматикой $S \to a$, $S \to bS$, $S \to cSS$?

2.3.4. Эквивалентны ли грамматика $T \to a$, $T \to Ta$, $T \to TT$, $T \to TTb$ и грамматика $R \to a$, $R \to RR$, $R \to aRb$?

2.3.5. Эквивалентны ли грамматика $T \to a$, $T \to TT$, $T \to TTb$ и грамматика $R \to a$, $R \to Ra$, $R \to RRb$?

2.3.6. Эквивалентны ли грамматика $T \to a$, $T \to Ta$, $T \to TT$, $T \to TTb$ и грамматика $R \to a$, $R \to Ra$, $R \to aRb$?

2.3.7. Эквивалентны ли грамматика $T \to bV$, $T \to VbT$, $T \to aTbTa$, $T \to \varepsilon$, $V \to TaaT$ и грамматика $R \to UaR$, $R \to aU$, $R \to \varepsilon$, $U \to bRaR$, $U \to aRbR$?

2.3.8. Эквивалентны ли грамматика $T \to bTaU$, $T \to VU$, $T \to UV$, $T \to \varepsilon$, $V \to UbT$, $U \to aT$ и грамматика $R \to bRaRaR$, $R \to aRbRa$, $R \to aRaRbR$, $R \to \varepsilon$?

2.3.9. Эквивалентны ли грамматика $T \to bTaU$, $T \to aTbU$, $T \to aUb$, $T \to \varepsilon$, $U \to TaT$ и грамматика $R \to bRaRaR$, $R \to aRbRa$, $R \to aRaRbR$, $R \to \varepsilon$?

2.3.10. Эквивалентны ли грамматика $T \to bTaaT$, $T \to aTbTaT$, $T \to aaTbT$, $T \to \varepsilon$ и грамматика $R \to bRaRa$, $R \to aRbRaR$, $R \to aRaRbR$, $R \to \varepsilon$?

2.3.11. Эквивалентны ли грамматика $S \to bSaT$, $S \to TbSa$, $S \to TTbS$, $S \to \varepsilon$, $T \to aS$ и грамматика $R \to bRUU$, $R \to UbU$, $R \to UUbR$, $R \to \varepsilon$, $U \to aR$?

2.3.12. Эквивалентны ли грамматика $S \to SaSbS$, $S \to c$ и грамматика $R \to cT$, $T \to acTbcT$, $T \to \varepsilon$?

2.3.13. Описать язык, порождаемый грамматикой $F \to ab$, $F \to aFb$, $F \to FF$.

2.3.14. Описать язык, порождаемый грамматикой $F \to a$, $F \to bF$, $F \to cFF$.

2.3.15. Описать язык, порождаемый грамматикой $F \to a$, $F \to bF$, $F \to cFF$, $F \to dFFF$.

2.3.16. Описать язык, порождаемый грамматикой $S \to \varepsilon$, $S \to aSSb$, $S \to bSSa$.

2.3.17. Описать язык, порождаемый грамматикой $F \to \varepsilon$, $F \to aFaFbF$.

2.3.18. Описать язык, порождаемый грамматикой $S \to \varepsilon$, $S \to aSSa$, $S \to bSSb$.

2.3.19. Найти контекстно-свободную грамматику с двумя правилами, эквивалентную грамматике $S \to \varepsilon$, $S \to F$, $F \to ab$, $F \to aFb$, $F \to FF$.

2.3.20. Найти однозначную контекстно-свободную грамматику, порождающую язык $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$.

2.3.21. Найти однозначную контекстно-свободную грамматику, порождающую язык $\{w \in \{a,b\}^* \mid |w|_a \geqslant |w|_b\}$.

2.3.22. Обозначим через L язык, порождаемый грамматикой $S \to aSbS$, $S \to \varepsilon$. Найти контекстно-свободную грамматику, порождающую язык $\{w \in L \mid \mathrm{Subw}(w) \cap (\{ba\} \cdot L \cdot \{ba\}) = \varnothing\}$.

3. Нормальные формы контекстно-свободных грамматик

3.1. Устранение бесполезных символов

[20, 8.], [2, 2.4.2], [27, 7.1.1, 7.1.2], [9, с. 427—429], [24, 4.2], [18, I.4.1, I.4.6], [5, с. 11—12], [7, 4.2], [4, с. 61—64], [22, с. 65—66], [13, 3.1.2], [14, с. 285—287], [17, 6.14]

Определение 1. Пусть дана порождающая грамматика $G = \langle N, \Sigma, P, S\rangle$. Символ $A \in N$ называется *полезным* (useful), если существуют такие слова $\alpha \in (N \cup \Sigma)^*$, $\beta \in (N \cup \Sigma)^*$ и $w \in \Sigma^*$, что $S \overset{*}{\Rightarrow} \alpha A \beta$ и $\alpha A \beta \overset{*}{\Rightarrow} w$. Символ $A \in N$ называется *бесполезным* (useless), если он не является полезным. Символ $A \in N$ называется *порождающим* (generating), если существует такое слово $w \in \Sigma^*$, что $A \overset{*}{\Rightarrow} w$. Символ $A \in N$ называется *достижимым* (reachable), если существуют такие слова $\alpha \in (N \cup \Sigma)^*$ и $\beta \in (N \cup \Sigma)^*$, что $S \overset{*}{\Rightarrow} \alpha A \beta$.

Лемма 2. *Пусть дана контекстно-свободная грамматика $G = \langle N, \Sigma, P, S\rangle$, в которой все символы из N являются порождающими. Пусть N' — множество всех достижимых символов грамматики G, а P' — множество тех правил из P, которые не содержат ни одного символа из множества $N – N'$. Тогда в контекстно-свободной грамматике $\langle N', \Sigma, P', S\rangle$ все символы из N' являются порождающими.*

Теорема 3. *Пусть дана контекстно-свободная грамматика $G = \langle N, \Sigma, P, S\rangle$ и $L(G) \neq \varnothing$. Тогда существуют такие множества $N' \subseteq N$ и $P' \subseteq P$, что в контекстно-свободной грамматике $\langle N', \Sigma, P', S\rangle$ нет бесполезных символов и она эквивалентна исходной грамматике.*

Доказательство. На первом этапе удалим все непорождающие символы (удалим также каждое правило, содержащее хотя бы один такой символ). На втором этапе из полученной грамматики удалим все недостижимые символы (и правила, их содержащие). Согласно лемме 2 на втором этапе ни один порождающий символ не может стать непорождающим. $\square$

Пример 4. Рассмотрим контекстно-свободную грамматику G с правилами $S \to UX$, $S \to VZ$, $T \to aa$, $T \to bb$, $U \to aUa$, $U \to bUb$, $V \to aTb$, $V \to bTa$, $W \to YZY$, $W \to aab$, $X \to Xa$, $X \to Xb$, $X \to \varepsilon$, $Y \to YY$, $Y \to aU$, $Y \to \varepsilon$, $Z \to W$, $Z \to b$. Удалив четыре правила, содержащие непорождающий символ U, получим грамматику G_1. В ней символ X является недостижимым. Удалив три правила, содержащие X, получим грамматику G_2 с правилами $S \to VZ$, $T \to aa$, $T \to bb$, $V \to aTb$, $V \to bTa$, $W \to YZY$, $W \to aab$, $Y \to YY$, $Y \to \varepsilon$, $Z \to W$, $Z \to b$. Очевидно, что $L(G) = L(G_2)$ и грамматика G_2 не содержит бесполезных символов.

Задачи

В задачах 3.1.1—3.1.10 надо найти контекстно-свободную грамматику без бесполезных символов, эквивалентную данной грамматике.

3.1.1. $S \to abS$, $S \to a$, $T \to aU$, $U \to baU$, $U \to \varepsilon$.

3.1.2. $R \to aRa$, $R \to b$, $R \to TT$, $T \to cR$, $V \to TT$, $V \to cR$.

3.1.3. $R \to aRaV$, $R \to b$, $R \to TT$, $T \to cTR$, $V \to TT$, $V \to cR$.

3.1.4. $S \to SRT$, $S \to c$, $R \to aRa$, $R \to b$, $T \to aT$.

3.1.5. $T \to aU$, $T \to bV$, $U \to aT$, $U \to bW$, $V \to aW$, $V \to bT$, $W \to aV$, $W \to bU$, $T \to \varepsilon$.

3.1.6. $J \to aaD$, $J \to FbF$, $D \to aaD$, $D \to FbH$, $H \to aDj$, $H \to Hb$, $F \to aJ$, $F \to Hb$, $F \to bFa$, $F \to \varepsilon$.

3.1.7. $S \to aaCM$, $S \to aaaKF$, $M \to aMb$, $M \to bMa$, $M \to \varepsilon$, $C \to aCa$, $C \to bCb$, $K \to bT$, $K \to aT$, $T \to bKa$, $T \to ab$.

3.1.8. $S \to aF$, $F \to J$, $F \to a$, $J \to F$, $J \to b$, $S \to aaCM$, $S \to aaaKF$, $M \to aMb$, $M \to bMa$, $M \to \varepsilon$, $C \to aCa$, $C \to bCb$, $K \to bT$, $K \to aT$, $T \to bKa$, $T \to ab$.

3.1.9. $S \to abTaaS$, $S \to abSbU$, $S \to bUU$, $T \to abTaaT$, $T \to abSbR$, $T \to bUR$, $T \to \varepsilon$, $R \to aSR$, $U \to aSU$, $U \to \varepsilon$.

3.1.10. $F \to aaFb$, $F \to abFa$, $F \to baFaba$, $F \to cFH$, $F \to c$, $H \to aaHb$, $H \to abHa$, $H \to baHaba$, $H \to cHF$.

3.2. Устранение эпсилон-правил

[20, 8.2], [6, 1.8], [2, 2.4.2], [27, 7.1.3], [18, I.2.4], [15, с. 295—296], [9, с. 429—431], [24, 2.3.2], [7, 4.2], [4, с. 71—76], [19, 7.4.3], [11, 7.2.6], [23, 6.2], [28, с. 216—217], [32, 2.1], [22, с. 35—37], [13, 3.1.2], [14, с. 288—289]

Теорема 1. *Пусть язык L является контекстно-свободным. Тогда язык $L - \{\varepsilon\}$ порождается некоторой контекстно-свободной грамматикой без ε-правил.*

Доказательство. Пусть дана контекстно-свободная грамматика $G = \langle N, \Sigma, P, S\rangle$, порождающая язык L. Проведём серию преобразований множества P.

Если для каких-то $A \in N$, $B \in N$, $\alpha \in (N \cup \Sigma)^*$ и $\beta \in (N \cup \Sigma)^*$ множество P содержит правила $B \to \alpha A \beta$ и $A \to \varepsilon$, но не содержит правила $B \to \alpha\beta$, то добавим это правило в P. Повторяем эту процедуру, пока возможно.

Теперь исключим из множества P все правила вида $A \to \varepsilon$. Полученная грамматика порождает язык $L - \{\varepsilon\}$. □

Пример 2. Рассмотрим язык L, порождаемый грамматикой $S \to \varepsilon$, $S \to aSbS$. Язык $L - \{\varepsilon\}$ порождается грамматикой $S \to aSbS$, $S \to abS$, $S \to aSb$, $S \to ab$.

Задачи

3.2.1. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $D \to Sc$, $S \to SS$, $S \to \varepsilon$, $S \to aSb$.

3.2.2. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $D \to DF$, $D \to ab$, $F \to \varepsilon$, $F \to FF$.

3.2.3. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $D \to cF$, $F \to FcF$, $F \to \varepsilon$, $F \to aDb$.

3.2.4. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $D \to Fc$, $F \to aaFb$, $F \to Fa$, $F \to \varepsilon$.

3.2.5. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $S \to AW$, $S \to baaBT$, $A \to a$, $A \to c$, $B \to b$, $B \to c$, $Z \to a$, $Z \to B$, $W \to ZW$, $W \to \varepsilon$,

$T \to Ubbaa$, $T \to Xbaa$, $T \to Xaa$, $T \to Ua$, $T \to U$, $U \to WB$, $U \to \varepsilon$, $X \to WA$, $X \to \varepsilon$.

3.2.6. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $S \to cEc$, $E \to \varepsilon$, $E \to aGbE$, $E \to bFaE$, $G \to aGbG$, $G \to \varepsilon$, $F \to bFaF$, $F \to \varepsilon$.

3.2.7. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $T \to ScS$, $S \to \varepsilon$, $S \to aSbbSa$, $S \to bSaaSb$, $S \to aSb$, $S \to bSa$.

3.2.8. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $T \to aE$, $E \to \varepsilon$, $E \to aE$, $E \to aEbE$, $E \to bEaE$.

3.2.9. Найти контекстно-свободную грамматику без ε-правил, эквивалентную грамматике $F \to HF$, $F \to aFbD$, $F \to bFaD$, $F \to cD$, $H \to aDb$, $H \to bDa$, $H \to \varepsilon$, $D \to HD$, $D \to \varepsilon$.

3.3. Нормальная форма Хомского

[20, 8.3], [23, 6.2], [2, 2.4.2, 2.4.3], [27, 7.1], [9, с. 431—437], [24, 4.1], [18, I.4.1, I.4.2], [7, 4.1], [25, с. 32], [15, с. 330], [32, 2.1], [13, 3.1.2], [22, с. 66—68], [19, 7.7.1], [8, с. 67]

Определение 1. *Грамматика в нормальной форме Хомского* (*грамматика в бинарной нормальной форме, квадратичная грамматика*, grammar in Chomsky normal form) — контекстно-свободная грамматика $\langle N, \Sigma, P, S\rangle$, в которой каждое правило имеет один из следующих трёх видов: $S \to \varepsilon$, $A \to a$, $A \to BC$, где $A \in N$, $B \in N - \{S\}$, $C \in N - \{S\}$, $a \in \Sigma$.

Пример 2. Грамматика $S \to RR$, $S \to AB$, $R \to RR$, $R \to AB$, $A \to a$, $B \to RB$, $B \to b$ является грамматикой в нормальной форме Хомского.

Теорема 3. *Каждая контекстно-свободная грамматика эквивалентна некоторой грамматике в нормальной форме Хомского.*

Доказательство. Пусть дана контекстно-свободная грамматика $G = \langle N, \Sigma, P, S\rangle$. Проведём ряд преобразований этой грамматики так, что порождаемый ею язык остаётся неизменным.

Если правая часть какого-нибудь правила содержит символ S, то заменим грамматику $\langle N, \Sigma, P, S\rangle$ на грамматику

$$\langle N \cup \{S_0\}, \Sigma, P \cup \{S_0 \to S\}, S_0\rangle,$$

где S_0 — новый символ, не принадлежащий множеству $N \cup \Sigma$.

Заменим во всех правилах каждый терминальный символ a на новый нетерминальный символ T_a и добавим к множеству P правила $T_a \to a$ для всех $a \in \Sigma$.

Устраним правила вида $A \to \alpha$, где $|\alpha| > 2$, заменив каждое из них на ряд более коротких правил (при этом добавляются новые нетерминальные символы).

Теперь устраним все правила вида $A \to \varepsilon$, где A не является начальным символом. Это можно сделать так же, как в доказательстве теоремы 1 на с. 34.

Если для каких-то $A \in N$, $B \in N$ и $\alpha \in (N \cup \Sigma)^*$ множество P содержит правила $A \to B$ и $B \to \alpha$, но не содержит правила $A \to \alpha$, то добавим это правило в P. Повторяем эту процедуру, пока возможно. После этого исключим из множества P все правила вида $A \to B$. $\square$

Пример 4. Грамматика $S \to \varepsilon$, $S \to aUbU$, $U \to S$, $U \to ba$ эквивалентна следующей грамматике в нормальной форме Хомского: $S_0 \to \varepsilon$, $S_0 \to AD$, $D \to UC$, $D \to BU$, $D \to b$, $C \to BU$, $C \to b$, $U \to BA$, $U \to AD$, $A \to a$, $B \to b$.

Теорема 5. *Если контекстно-свободный язык не содержит пустого слова, то он порождается некоторой грамматикой, в которой каждое правило имеет один из следующих двух видов*: $A \to a$, $A \to BC$, *где* $A \in N$, $B \in N - \{S\}$, $C \in N - \{S\}$, $a \in \Sigma$.

Задачи

3.3.1. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $F \to ab$, $F \to aFb$, $F \to FF$.

3.3.2. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $F \to a$, $F \to bF$, $F \to cFF$.

3.3.3. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $S \to RS$, $S \to R$, $R \to aSb$, $R \to cSd$, $R \to ab$, $R \to cd$.

3.3.4. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $S \to Raabb$, $aab \to abbaa$, $Ra \to Raa$, $R \to b$.

3.3.5. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $J \to Jb$, $Jbb \to bJc$, $bJb \to c$.

3.3.6. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $S \to F$, $S \to T$, $F \to \varepsilon$, $F \to aJ$, $F \to bF$, $J \to aF$, $J \to bJ$, $T \to a$, $T \to bT$, $T \to Tb$, $T \to TaT$.

3.3.7. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $S \to F$, $S \to T$, $F \to aaF$, $F \to bF$, $F \to \varepsilon$, $T \to a$, $T \to bT$, $T \to Tb$, $T \to TaT$.

3.3.8. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $S \to AW$, $S \to baaBT$, $A \to a$, $A \to c$, $B \to b$, $B \to c$, $Z \to a$, $Z \to B$, $W \to ZW$, $W \to \varepsilon$, $T \to Ubbaa$, $T \to Xbaa$, $T \to Xaa$, $T \to Ua$, $T \to U$, $U \to WB$, $U \to \varepsilon$, $X \to WA$, $X \to \varepsilon$.

3.3.9. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $R \to aEa$, $E \to \varepsilon$, $E \to aaE$, $E \to bbE$, $E \to abEbaE$, $E \to baEabE$.

3.4. Нормальная форма Грейбах

[20, 8.4], [23, 6.2], [2, 2.4.4, 2.4.5], [18, I.4.3], [9, с. 437—440], [7, 6.2], [4, с. 121—124], [22, с. 71—76], [13, 3.1.2], [25, с. 32]

Определение 1. *Грамматика в нормальной форме Грейбах* (grammar in Greibach normal form) — контекстно-свободная грамматика $\langle N, \Sigma, P, S\rangle$, в которой каждое правило имеет один из следующих четырёх видов: $A \to \varepsilon$, $A \to a$, $A \to aB$, $A \to aBC$, где $A \in N$, $B \in N$, $C \in N$, $a \in \Sigma$.

Пример 2. Грамматика $S \to aST$, $S \to aT$, $T \to bS$, $T \to b$ является грамматикой в нормальной форме Грейбах.

Замечание 3. Некоторые авторы разрешают в грамматиках в нормальной форме Грейбах использовать также правила вида $A \to a\beta$, где $A \in N$, $a \in \Sigma$, $\beta \in N^*$ (в определении 1 такие правила разрешены, только если $|\beta| \leqslant 2$).

Теорема 4. *Каждая контекстно-свободная грамматика эквивалентна некоторой грамматике в нормальной форме Грейбах.*

Доказательство теоремы 4 можно найти в [20, 8.4].

Теорема 5. *Пусть язык L контекстно-свободный. Тогда язык $L - \{\varepsilon\}$ порождается некоторой грамматикой в нормальной форме Грейбах без ε-правил.*

Задачи

3.4.1. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $F \to ab$, $F \to aFb$, $F \to FF$.

3.4.2. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $S \to AB$, $B \to AB$, $A \to BB$, $A \to a$, $B \to b$.

3.4.3. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $K \to Fb$, $F \to \varepsilon$, $F \to aFbF$.

3.4.4. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $K \to FJ$, $J \to aa$, $J \to abJ$, $J \to bJJ$, $F \to aK$, $F \to bJ$.

3.4.5. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $S \to a$, $S \to Sb$, $S \to SSc$.

3.4.6. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $S \to aT$, $T \to aTTa$, $T \to b$.

3.4.7. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $R \to aEa$, $E \to \varepsilon$, $E \to aaE$, $E \to bbE$, $E \to abEbaE$, $E \to baEabE$.

3.4.8. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $F \to \varepsilon$, $F \to FbFaF$.

3.4.9. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $F \to \varepsilon$, $F \to bFFa$.

3.4.10. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $F \to \varepsilon$, $F \to FF$, $F \to bFFa$.

3.4.11. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $R \to TR$, $R \to a$, $T \to R$, $T \to bR$.

3.4.12. Найти контекстно-свободную грамматику в нормальной форме Грейбах, эквивалентную грамматике $\langle N, \Sigma, P, S\rangle$, где $\Sigma = \{\mathrm{a}, \mathrm{the}, \mathrm{that}, \mathrm{table}, \mathrm{face}, \mathrm{little}, \mathrm{red}\}$, $N = \{\mathrm{Det}, \mathrm{N}, \mathrm{Adj}, \mathrm{NP}, \check{\mathrm{N}}\}$, $S = \mathrm{NP}$,

$$P = \{\mathrm{Det} \to \mathrm{a}, \mathrm{Det} \to \mathrm{the}, \mathrm{Det} \to \mathrm{that}, \mathrm{N} \to \mathrm{table}, \mathrm{N} \to \mathrm{face},$$
$$\mathrm{Adj} \to \mathrm{little}, \mathrm{Adj} \to \mathrm{red}, \mathrm{NP} \to \mathrm{Det}\ \check{\mathrm{N}}, \mathrm{NP} \to \check{\mathrm{N}},$$
$$\check{\mathrm{N}} \to \mathrm{Adj}\ \check{\mathrm{N}}, \check{\mathrm{N}} \to \mathrm{N}\}.$$

4. Основные свойства контекстно-свободных языков

4.1. Лемма о разрастании для контекстно-свободных языков

[20, 9.1], [32, 2.3], [6, 3.1], [27, 7.2], [15, с. 299—300], [2, 2.6.1], [7, 4.3, 4.1], [9, с. 425—426], [24, 4.1], [25, с. 33—34], [31, 3.5], [11, 7.4.2], [22, с. 68—71], [14, с. 279—281]

Лемма 1 (pumping lemma, лемма о разрастании, лемма о накачке, лемма-насос)**.** *Пусть L — контекстно-свободный язык над алфавитом Σ. Тогда найдётся такое положительное целое число p, что для любого слова $w \in L$ длины не меньше p можно подобрать слова $u, v, x, y, z \in \Sigma^*$, для которых верно $uvxyz = w$, $vy \neq \varepsilon$ (то есть $v \neq \varepsilon$ или $y \neq \varepsilon$), $|vxy| \leqslant p$ и $uv^ixy^iz \in L$ для всех $i \in \mathbb{N}$.*

Доказательство можно найти в [20, 9.1].

Пример 2. Рассмотрим язык $L = \{a^nb^nc^n \mid n \geqslant 0\}$ над алфавитом $\{a, b, c\}$. Утверждение леммы 1 не выполняется ни для какого натурального числа p. Действительно, если $uvxyz = a^pb^pc^p$, $|vy| > 0$ и $|vxy| \leqslant p$, то $|vy|_a = 0$ или $|vy|_c = 0$. Следовательно, $|uvvxyyz|_a = p$ или $|uvvxyyz|_c = p$. Так как $|uvvxyyz| > 3p$, то $uvvxyyz \notin L$. Из этого можно заключить, что язык L не является контекстно-свободным.

Теорема 3. *Каждый контекстно-свободный язык над однобуквенным алфавитом является автоматным.*

Доказательство можно найти в [20, 9.1].

Задачи

В задачах 4.1.1—4.1.59 надо выяснить, является ли данный язык контекстно-свободным.

4.1.1. $\{a^nb^nc^n \mid n \geqslant 1\}$.

4.1.2. $\{v_1cv_1 \mid v_1 \in \{a, b\}^*\}$.

4.1.3. $\{v_1v_1 \mid v_1 \in \{a, b\}^*\}$.

4.1.4. $\{v_1v_1v_1 \mid v_1 \in \{a, b\}^*\}$.

4.1.5. $\{v_1v_2v_2v_3 \mid v_1 \in \{a,b\}^+,\ v_2 \in \{a,b\}^+,\ v_3 \in \{a,b\}^+\}$.

4.1.6. $\{a^ib^jc^k \mid 0 \leqslant i \leqslant j \leqslant k\}$.

4.1.7. $\{c^mdc^mdc^m \mid m \geqslant 0\}$.

4.1.8. $\{c^mdc^{2m}dc^{3m} \mid m \geqslant 0\}$.

4.1.9. $\{a^mb^n \mid (m+n) > 0,\ (m+n) \vdots 2\}$.

4.1.10. $\{a^kb^la^mb^n \mid k = l \text{ или } m = n\}$.

4.1.11. $\{a^kb^la^mb^n \mid k = m \text{ или } l = n\}$.

4.1.12. $\{a^kb^mb^na^m \mid n \geqslant k \geqslant 0,\ m \geqslant 0\}$.

4.1.13. $\{a^kb^la^mb^n \mid k + l = m + n\}$.

4.1.14. $\{ba^kba^lba^mb \mid k \neq 1 \text{ или } l \neq 2 \text{ или } m \neq 3\}$.

4.1.15. $\{a^kb^ma^nb^m \mid n \geqslant k \geqslant 0,\ m \geqslant 0\}$.

4.1.16. $\{a^kb^la^mb^n \mid l \geqslant n \geqslant 0,\ k \geqslant m \geqslant 0\}$.

4.1.17. $\{a^kb^la^mb^n \mid l \geqslant k \geqslant 0,\ n \geqslant m \geqslant 0\}$.

4.1.18. $\{a^kb^la^mb^n \mid k < l,\ m < n\}$.

4.1.19. $\{a^kb^la^mb^n \mid k < m,\ l < n\}$.

4.1.20. $\{a^kb^la^mb^n \mid k < n,\ l < m\}$.

4.1.21. $\{a^mb^nc^kd^l \mid m \leqslant k,\ m + n = k + l\}$.

4.1.22. $\{a^iba^jba^kb^ka^jb^la^lb^i \mid i \geqslant 3,\ j \geqslant 2,\ k \geqslant 1,\ l \geqslant 0\}$.

4.1.23. $\{(ab)^ka^na^k(ab)^n \mid k \geqslant 0,\ n \geqslant 0\}$.

4.1.24. $\{a^kb^{l+2m}a^kb^l \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\} \cup$ $\cup \{a^kb^la^{k+2m}b^l \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\}$.

4.1.25. $\{a^kb^{l+2m}a^kb^l \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\} \cup$ $\cup \{a^kb^la^kb^{l+2m} \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\}$.

4.1.26. $\{a^nb^na^n \mid n \geqslant 0\} \cup \{a^{2k}b^la^m \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\} \cup$ $\cup \{a^kb^{2l+1}a^m \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\}$.

4.1.27. $\{a^nb^na^n \mid n \geqslant 0\} \cup \{a^kb^la^m \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 52\}$.

4.1.28. $\{a^{k+m}b^ma^kb^m \mid k \geqslant 0,\ m \geqslant 0\}$.

4.1.29. $\{a^{k+m}b^ma^kb^n \mid k \geqslant 0,\ m \geqslant 0,\ n = \max(k-m, 0)\}$.

4.1.30. $\{a^kb^{k+m}a^kb^m \mid k \geqslant 0,\ m \geqslant 0\}$.

4.1.31. $\{a^m \mid m \text{ простое}\}$.

4.1.32. $\{a^mb^n \mid m \leqslant n,\ m \text{ простое}\}$.

4.1.33. $\{a^{n^2} \mid n \geqslant 1\}$.

4.1.34. $\{a^nb^na^n \mid n \geqslant 0\} \cup \{a^{2k}b^la^m \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\}$.

4.1.35. $\{a^nb^na^n \mid n \geqslant 0\} \cup \{a^kb^la^m \mid k \geqslant 0,\ l \geqslant 0,\ m \leqslant 52\}$.

4.1.36. $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c\}$.

4.1.37. $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$.

4.1.38. $\{w \in \{a,b\}^* \mid |w|_a = |w|_b + 1\}$.

4.1.39. $\{w \in \{a,b\}^* \mid |w|_a > |w|_b + 1\}$.

4.1.40. $\{a^n b^n a^n \mid n \geqslant 1\} \cup \{w \in \{a,b\}^* \mid |w|_a > |w|_b + 2\}$.

4.1.41. $\{w \in \{a,b,c\}^* \mid |w|_a \neq |w|_b + |w|_c,\ |w|_a + 2|w|_b > 4,\ (|w|_a + |w|_c) \,\vdots\, 4\}$.

4.1.42. $\{a^n b v_1 b a^m \mid v_1 \in \{a,b\}^*,\ |v_1|_a = m + n\}$.

4.1.43. $\{v_1 v_2 v_3 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{c,d\}^*,\ v_3 \in \{a,b\}^*,\ |v_1|_a = |v_3|_b,\ |v_2|_c = |v_2|_d\}$.

4.1.44. $\{v_1 v_2 v_3 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{c,d\}^*,\ v_3 \in \{a,b\}^*,\ |v_1|_a = |v_2|_c,\ |v_2|_d = |v_3|_b\}$.

4.1.45. $\{v_1 v_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ v_1 = v_1^{\mathrm{R}},\ v_2 = v_2^{\mathrm{R}}\} - \{\varepsilon\}$.

4.1.46. $\{c^n d c^m d c^{nm} \mid n \geqslant 0,\ m \geqslant 0\}$.

4.1.47. $\{baba^2ba^3b \ldots ba^{n-1}ba^nb \mid n \geqslant 1\}$.

4.1.48. $\{ba^{n_1}ba^{n_2}b \ldots a^{n_p}b \mid \exists j\, n_j \neq j\}$.

4.1.49. $\{v_2 c v_1 v_2 v_3 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ v_3 \in \{a,b\}^*\}$.

4.1.50. $\{v_2 c v_1 v_2^{\mathrm{R}} v_3 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ v_3 \in \{a,b\}^*\}$.

4.1.51. $\{v_1 v_2 v_2^{\mathrm{R}} v_1 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*\}$.

4.1.52. $\{a^n b^n a^n b^n \mid n \geqslant 0\} \cup \{a^{2k} b^{2k} a^m b^m \mid k \geqslant 0,\ m \geqslant 0\} \cup \cup \{a^k b^{2l} b a a^{2l} b^k \mid k \geqslant 0,\ m \geqslant 0\}$.

4.1.53. $\{a^k b^{l+2m} a^k b^l \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\} \cup \cup \{a^k b^l a^k b^{l+2m} \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\}$.

4.1.54. $\{a^k b^m c^n \mid k < \max(m,n)\}$.

4.1.55. $\{a^k b^m c^n \mid k > \max(m,n)\}$.

4.1.56. $\{a^k b^m c^n \mid m > \max(k,n)\}$.

4.1.57. $\{a^k b^m c^n \mid k = \max(m,n)\}$.

4.1.58. $\{a^m b^n \mid 7m < n < 10 + m^2\}$.

4.1.59. $\{a^m b^n \mid 10 + m^2 < n < 7m\}$.

В задачах 4.1.60—4.1.69 надо выяснить, какому классу данный язык принадлежит.

4.1.60. $\{b^n a^n c^n \mid n \geqslant 1\}$.

4.1.61. $\{a^{n^2} \mid n \geqslant 1\}$.

4.1.62. $\{a^n b^n c^m \mid n \geqslant 1,\ m \geqslant 1\}$.

4.1.63. $\{a^n b^m c^{n+m} \mid n \geqslant 1,\ m \geqslant 1\}$.

4.1.64. $\{a^n b^m c^{2n+3m} \mid n \geqslant 1,\ m \geqslant 1\}$.

4.1.65. $\{a^n b^m c^{nm} \mid n \geqslant 1,\ m \geqslant 1\}$.

4.1.66. $\{a^n b^m c^k \mid n \geqslant 1,\ m \geqslant 1,\ k \geqslant 1\}$.

4.1.67. $\{a(cd)^n a \mid n \geqslant 0\} \cup \{b(cd)^n b \mid n \geqslant 0\}$.

4.1.68. $\{a^n b^m c^{nm} \mid n \geqslant 1,\ m \geqslant 1\} \cap \{a^{nm} b^n c^m \mid n \geqslant 1,\ m \geqslant 1\}$.

4.1.69. $\{(abba)^{n^2} \mid n \geqslant 1\} \cup \{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$.

4.1.70. Является ли контекстно-свободным язык, порождаемый грамматикой $F \to d$, $F \to Gbdc$, $F \to aFJ$, $cJ \to cbaFbFc$, $dJ \to FbFc$, $G \to ad$, $G \to aGbFc$?

4.1.71. Является ли контекстно-свободным язык, порождаемый грамматикой $F \to HAK$, $H \to bK$, $K \to GKG$, $K \to \varepsilon$, $AG \to G$, $A \to a$, $G \to c$?

4.1.72. Существуют ли такие контекстно-свободные языки $L_1 \subseteq \{a,b\}^*$ и $L_2 \subseteq \{b,c\}^*$, что язык $L_1 \cap L_2$ не является автоматным?

4.1.73. Существуют ли такие контекстно-свободные языки $L_1 \subseteq \{a,b\}^*$ и $L_2 \subseteq \{b,c\}^*$, что язык $L_1 - L_2$ не является контекстно-свободным?

4.1.74. Существуют ли такие контекстно-свободные языки $L_1 \subseteq \{a,b\}^*$ и $L_2 \subseteq \{b,c\}^*$, что язык $L_1 - L_2$ не является автоматным?

4.2. Лемма о разрастании для линейных языков

[20, 9.2], [2, с. 191, 237]

Лемма 1. *Пусть L — линейный язык над алфавитом Σ. Тогда найдётся такое положительное целое число p, что для любого слова $w \in L$ длины не меньше p можно подобрать слова $u, v, x, y, z \in \Sigma^*$, для которых верно $uvxyz = w$, $vy \neq \varepsilon$ (то есть $v \neq \varepsilon$ или $y \neq \varepsilon$), $|uv| + |yz| \leqslant p$ и $uv^ixy^iz \in L$ для всех $i \in \mathbb{N}$.*

Доказательство можно найти в [20, 9.2].

Задачи

В задачах 4.2.1—4.2.23 надо выяснить, является ли данный язык линейным.

4.2.1. $\{a^m b^m a^n b^n \mid m \geqslant 0,\ n \geqslant 0\}$.

4.2.2. $\{a^m b^n \mid m < n\}$.

4.2.3. $\{a^n b^n c^m d^m \mid n \geqslant 1,\ m \geqslant 1\}$.

4.2.4. $\{a^n b^n (cd)^m \mid n \geqslant 1,\ m \geqslant 1\}$.

4.2.5. $\{a^k b^l c^m d^n \mid k < n,\ l > m\}$.

4.2.6. $\{a^k b^l c^m a^n b^n c^k \mid k \geqslant 1,\ l < m,\ n \geqslant 3\}$.

4.2.7. $\{a(cb)^n b(bc)^n \mid n \geqslant 0\}$.

4.2.8. $\{a^k b^l c^m d^n \mid k < l,\ n < m\}$.

4.2.9. $\{a^k b^l c^m d^n \mid l < k,\ m < n\}$.

4.2.10. $\{a^k b^l c^m d^n \mid k < m,\ l < n\}$.

4.2.11. $\{b^m a(bc)^m \mid m \geqslant 0\}$.

4.2.12. $\{a^k b^m a^n \mid k \geqslant 1,\ n \geqslant m \geqslant 0\}$.

4.2.13. $\{a^k b^m a^n \mid k \neq n$ или $m = 0\}$.

4.2.14. $\{a^k c^m b^n \mid k \geqslant 0,\ n \geqslant 0,\ m = k+n+1\}$.

4.2.15. $\{a^k c^m b^n \mid k \geqslant 0,\ m \geqslant 0,\ n = k+m\}$.

4.2.16. $\{v_1 v_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ |v_1|_b = |v_2|_a\}$.

4.2.17. $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$.

4.2.18. $\{w \in \{a,b\}^* \mid |w|_a > |w|_b + 2\}$.

4.2.19. $\{w \in \{a,b\}^+ \mid w = w^{\mathrm{R}}\}$.

4.2.20. $\{v_1 v_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ |v_1|_a = |v_2|_a\}$.

4.2.21. $\{v_1 v_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ |v_1|_a = |v_1|_b,\ |v_2| < 4\}$.

4.2.22. $\{v_1 v_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ |v_1|_a \leqslant |v_2|_b,\ |v_1|_b \leqslant |v_2|_a\}$.

4.2.23. $\{v_1 v_1 \mid v_1 \in \{a,b\}^*,\ |v_1| < 4\}$.

4.3. Свойства замкнутости класса линейных языков

[20, 9.3], [2, 2.6.4]

Теорема 1. *Класс линейных языков замкнут относительно объединения и обращения.*

Пример 2. Рассмотрим алфавит $\Sigma = \{a, b, c\}$. Язык

$$L = \Sigma^* - \{a^n b^n c^n \mid n \geqslant 0\}$$

является линейным, поскольку

$$L = L_1 \cup L_2 \cup (\Sigma^* - L_3),$$

где языки

$$L_1 = \{a^m b^n c^k \mid m \neq n,\ k \geqslant 0\}, \quad L_2 = \{a^m b^n c^k \mid n \neq k,\ m \geqslant 0\}$$

являются линейными, а язык

$$L_3 = \{a^m b^n c^k \mid m \geqslant 0,\ n \geqslant 0,\ k \geqslant 0\}$$

является автоматным, и можно применить теоремы о связи классов автоматных, праволинейных и линейных языков, а также свойства замкнутости этих классов.

Теорема 3. *Если язык L линейный, то L^{R} тоже линейный язык.*

Теорема 4. *Если язык L_1 автоматный и язык L_2 линейный, то языки $L_1 \cdot L_2$ и $L_2 \cdot L_1$ — линейные языки.*

Задачи

В задачах 4.3.1—4.3.37 надо выяснить, является ли данный язык линейным.

4.3.1. $\{w \in \{a,b\}^* \mid |w|_a \,\vdots\, 3\} \cup \{(ba)^m a^{2n} \mid m \geqslant 0,\ n \geqslant 0\}$.

4.3.2. $\{w \in \{a,b\}^* \mid |w|_a \,\vdots\, 3\} \cdot \{(ba)^m a^{2m} \mid m \geqslant 0\}$.

4.3.3. $\{a^m b^n \mid m \neq n\}$.

4.3.4. $\{a^m b^n c^k \mid m < n,\ k \geqslant 0\}$.

4.3.5. $\{a^m b^n c \mid m \neq n\}$.

4.3.6. $\{a^m b^n a b^{2n+1} \mid m \geqslant 0,\ n \geqslant 0\}$.

4.3.7. $\{a^n b^n c^m \mid n \geqslant 1,\ m \geqslant 1\}$.

4.3.8. $\{a^n b^n c^m \mid n \geqslant 1,\ m \geqslant 1\} \cup \{a^n b^m c^m \mid n \geqslant 1,\ m \geqslant 1\}$.

4.3.9. $\{a^i b^j a^k \mid i \neq j$ или $j \neq k\}$.

4.3.10. $\{a^m b^n c^k d^l \mid m + n = k + l\}$.

4.3.11. $\{a^m b^n c^k d^l \mid n = l$ или $m \leqslant k$ или $m + n = k + l\}$.

4.3.12. $\{a^m b^n c^k d^l \mid n \leqslant k,\ m + n = k + l\}$.

4.3.13. $\{a^{2m} u \mid u \in \{a,b\}^*,\ m \geqslant 0,\ |u|_b = m\}$.

4.3.14. $\{a^m u \mid u \in \{a,b\}^*,\ m \geqslant 0,\ |u|_b = 2m\}$.

4.3.15. $\{a^n b^m c^{2k} d^k e^n \mid n \geqslant 1,\ m \geqslant 1,\ k \geqslant 1\}$.

4.3.16. $\{a^k c (ac)^k a^m \mid k \geqslant 0,\ m \geqslant 0\}$.

4.3.17. $\{(bc)^m b^n c b^m \mid m \geqslant 0,\ n \geqslant 0\}$.

4.3.18. $\{c^k d^n c (cd)^n \mid k \geqslant 0,\ n \geqslant 0\}$.

4.3.19. $\{c^n d c^m d c^{n+m} \mid n \geqslant 0,\ m \geqslant 0\}$.

4.3.20. $\{b u u^{\mathrm{R}} b \mid u \in \{a,b\}^*\}$.

4.3.21. $\{a u \mid u \in \{a,b\}^*,\ u = u^{\mathrm{R}}\}$.

4.3.22. $\{v_1 c v_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ |v_1| = |v_2|\}$.

4.3.23. $\{v_1av_2b \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ |v_1| = |v_2|\}$.

4.3.24. $\{v_1cv_2cv_3 \mid v_1, v_2, v_3 \in \{a,b\}^*,\ v_1 = v_2^{\mathrm{R}}$ или $v_1 = v_3^{\mathrm{R}}$ или $v_2 = v_3^{\mathrm{R}}\}$.

4.3.25. $\{v_1av_2cv_3bv_4 \mid v_1,v_2,v_3,v_4 \in \{a,b\}^*, |v_1| = |v_4|, v_2 = v_3^{\mathrm{R}}\}$.

4.3.26. $\{v_1av_2cv_3bv_4 \mid v_1,v_2,v_3,v_4 \in \{a,b\}^*, v_1 = v_4^{\mathrm{R}}, |v_2| = |v_3|\}$.

4.3.27. $\{v_1cv_2 \mid v_1, v_2 \in \{a,b\}^*,\ |v_1| = |v_2|,\ v_1 \neq v_2^{\mathrm{R}}\}$.

4.3.28. $\{v_1av_2cv_3bv_4 \mid v_1, v_2, v_3, v_4 \in \{a,b\}^*,\ |v_1| = |v_3|\}$.

4.3.29. $\{a,b\}^* - \{a^nb^na^n \mid n \geqslant 0\}$.

4.3.30. $\{a,b,c\}^* - \{a^nb^nc^n \mid n \geqslant 0\}$.

4.3.31. $\{c,d\}^+ - \{c^ndc^ndc^n \mid n \geqslant 0\}$.

4.3.32. $\{v_1cv_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ v_1 \neq v_2^{\mathrm{R}}\}$.

4.3.33. $\{a,b,c\}^* - \{ucu^{\mathrm{R}} \mid u \in \{a,b\}^*\}$.

4.3.34. $\{v_1cv_2 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ v_1 \neq v_2\}$.

4.3.35. $\{a,b,c\}^* - \{ucu \mid u \in \{a,b\}^*\}$.

4.3.36. $\{w_1cw_2c \ldots cw_kccw_j^{\mathrm{R}} \mid 1 \leqslant j \leqslant k,\ \forall i\, w_i \in \{a,b\}^*\}$.

4.3.37. $\{a^nb^mc^k \mid n \geqslant 0,\ m \geqslant 1,\ k = \max(n - 2m, 0)\}$.

4.3.38. Найти линейную грамматику, порождающую язык $\{a,b,c\}^* - L_1$, где L_1 порождается грамматикой $F \to baFaaa$, $F \to baaFb$, $F \to bacaaa$, $F \to baacb$.

4.3.39. Найти линейную грамматику, порождающую язык $\{a,b,c\}^* - L_1$, где L_1 порождается грамматикой $F \to baFa$, $F \to baaFaab$, $F \to baaaFbaa$, $F \to baca$, $F \to baacaab$, $F \to baaacbaa$.

4.4. Свойства замкнутости класса контекстно-свободных языков

[20, 9.4], [11, 7.3.1—7.3.3], [15, с. 300—301], [27, 7.3.1—7.3.3], [2, 2.6.2], [6, 1.7], [18, I.9.1], [7, 4.3], [9, с. 423—424], [31, 3.5]

Теорема 1. *Класс контекстно-свободных языков замкнут относительно итерации, конкатенации, объединения и обращения.*

Задачи

4.4.1. Является ли контекстно-свободным язык $\{w \in \{a,b\}^* \mid |w|_a \neq |w|_b\}$?

4.4.2. Является ли контекстно-свободным язык $\{w \in \{a,b,c\}^* \mid |w|_a \neq |w|_b$ или $|w|_b \neq |w|_c\}$?

4.4.3. Найти контекстно-свободную грамматику для языка $L_1 \cup L_2$, где L_1 порождается грамматикой $F \to a$, $F \to bF$, $F \to cFF$, а язык L_2 порождается грамматикой $F \to aM$, $F \to cM$, $M \to aF$, $M \to bF$, $M \to \varepsilon$.

4.4.4. Найти контекстно-свободную грамматику с тремя правилами для языка $L_1 \cup L_2$, где L_1 порождается грамматикой $F \to aFa$, $F \to b$, а язык L_2 порождается грамматикой $F \to aMa$, $M \to aMa$, $M \to c$.

4.4.5. Найти самую короткую контекстно-свободную грамматику для языка $L_1 \cup L_2$, где L_1 порождается грамматикой $F \to aFa$, $F \to b$, а язык L_2 порождается грамматикой $F \to bFb$, $F \to a$.

4.4.6. Найти контекстно-свободную грамматику с тремя правилами для языка $\{a^k b^k a^m b^m a^n b^n \mid k \geqslant 2,\ m \geqslant 0,\ n \geqslant 1\}$.

4.5. Пересечение и дополнение контекстно-свободных языков

[20, 9.5], [6, 3.2], [2, 2.6.2], [27, 7.3.4], [18, I.9.1], [7, 4.3], [15, с. 300—301], [31, 3.5], [11, 7.3.4, 7.3.5], [9, с. 424], [22, с. 87—88]

Теорема 1. *Класс контекстно-свободных языков не замкнут ни относительно пересечения, ни относительно дополнения.*

Задачи

4.5.1. Является ли контекстно-свободным язык $\{a\}^* - \{a^{n^2} \mid n \geqslant 0\}$?

4.5.2. Является ли контекстно-свободным язык $\{a,b\}^* - \{uu \mid u \in \{a,b\}^*\}$?

4.5.3. Существует ли такой линейный язык L над алфавитом $\{a,b\}$, что язык $L^{\mathrm{R}} \cap L$ не является контекстно-свободным?

4.5.4. Пусть $\Sigma = \{a, b, c\}$. Рассмотрим язык L_1, порождаемый грамматикой $F \to baFa$, $F \to baaFaab$, $F \to baaaFbaa$, $F \to baca$, $F \to baacaab$, $F \to baaacbaa$, и язык L_2, порождаемый грамматикой $J \to baJaa$, $J \to baaJbb$, $J \to baaaJa$, $J \to bacaa$, $J \to baacbb$, $J \to baaaca$. Верно ли, что $L_1 \cap L_2 = \varnothing$?

4.5.5. Пусть $\Sigma = \{a, b, c, d_1, d_2, d_3\}$. Рассмотрим язык L_1, порождаемый грамматикой $F \to d_1Fa$, $F \to d_2Faab$, $F \to d_3Fbaa$, $F \to c$, и язык L_2, порождаемый грамматикой $J \to d_1Jaa$, $J \to d_2Jbb$, $J \to d_3Ja$, $J \to c$. Является ли язык $L_1 \cap L_2$ конечным?

4.5.6. Пусть $\Sigma = \{a, b, c, d_1, d_2, d_3\}$. Рассмотрим язык L_1, порождаемый грамматикой $F \to d_1Fab$, $F \to d_2Fb$, $F \to d_3Faba$, $F \to c$, и язык L_2, порождаемый грамматикой $J \to d_1Jaba$, $J \to d_2Ja$, $J \to d_3Jb$, $J \to c$. Является ли язык $L_1 \cap L_2$ контекстно-свободным?

4.6. Пересечение контекстно-свободного языка с автоматным языком

[20, 9.6], [6, 3.2], [2, 2.6.2], [27, 7.3.4], [18, I.9.1], [7, 5.2], [15, с. 301], [31, 3.5], [11, 10.6.3], [26, 1.12], [22, с. 87—88]

Теорема 1. *Если L_1 — контекстно-свободный язык и L_2 — автоматный язык, то язык $L_1 \cap L_2$ является контекстно-свободным.*

Доказательство. Пусть $\langle N, \Sigma, P, S \rangle$ — контекстно-свободная грамматика, порождающая язык L_1. Без ограничения общности можно считать, что множество P содержит только правила вида $A \to a$ и $A \to \alpha$, где $A \in N$, $a \in \Sigma$ и $\alpha \in N^*$ (см. теорему 3 на с. 35). Пусть $\langle Q, \Sigma, \Delta, I, F \rangle$ — конечный автомат, распознающий язык L_2. Без ограничения общности можно считать, что для каждого перехода $\langle p, x, q \rangle \in \Delta$ выполняется равенство $|x| = 1$.

Построим контекстно-свободную грамматику $\langle \overline{N}, \Sigma, \overline{P}, \overline{S} \rangle$, порождающую язык $L_1 \cap L_2$. Положим

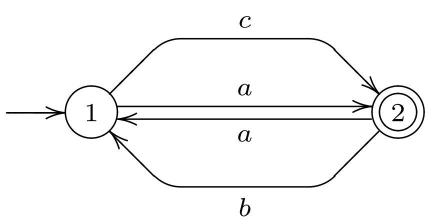
$$\overline{N} = \{\overline{S}\} \cup (Q \times N \times Q),$$
$$\overline{P} = \{\overline{S} \to \langle p, S, q \rangle \mid p \in I,\ q \in F\} \cup$$

$$\cup \{\langle p, A, q\rangle \to a \mid \langle p, a, q\rangle \in \Delta,\ (A \to a) \in P\} \cup$$
$$\cup \{\langle p_0, A, p_n\rangle \to \langle p_0, B_1, p_1\rangle \ldots \langle p_{n-1}, B_n, p_n\rangle \mid$$
$$(A \to B_1 \ldots B_n) \in P,\ p_0 \in Q, \ldots,\ p_n \in Q\},$$

где $\overline{S}$ — новый символ (не принадлежащий множеству $Q \times N \times Q$). $\square$

Пример 2. Пусть $\Sigma = \{a, b, c\}$. Рассмотрим контекстно-свободный язык L_1, порождаемый грамматикой $S \to a$, $S \to bS$, $S \to cSS$, и автоматный язык L_2, распознаваемый конечным автоматом $M = \langle Q, \Sigma, \Delta, I, F\rangle$, где $Q = \{1, 2\}$, $I = \{1\}$, $F = \{2\}$,

$$\Delta = \{\langle 1, a, 2\rangle,\ \langle 1, c, 2\rangle,\ \langle 2, a, 1\rangle,\ \langle 2, b, 1\rangle\}.$$

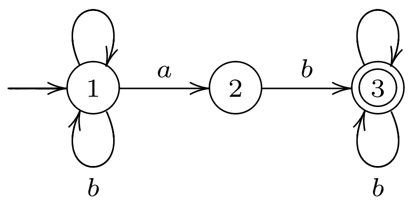


Тогда язык $L_1 \cap L_2$ порождается контекстно-свободной грамматикой $\overline{S} \to S_{12}$, $S_{12} \to a$, $S_{21} \to a$, $S_{21} \to bS_{11}$, $S_{22} \to bS_{12}$, $S_{11} \to cS_{21}S_{11}$, $S_{11} \to cS_{22}S_{21}$, $S_{12} \to cS_{21}S_{12}$, $S_{12} \to cS_{22}S_{22}$. Здесь S_{11}, S_{12}, S_{21} и S_{22} соответствуют символам $\langle 1, S, 1\rangle$, $\langle 1, S, 2\rangle$, $\langle 2, S, 1\rangle$ и $\langle 2, S, 2\rangle$ из доказательства теоремы 1.

Теорема 3. *Если L_1 — линейный язык и L_2 — автоматный язык, то язык $L_1 \cap L_2$ является линейным.*

Пример 4. Пусть $\Sigma = \{a, b\}$. Рассмотрим линейный язык L_1, порождаемый грамматикой $S \to aaSaa$, $S \to bSb$, $S \to \varepsilon$, и автоматный язык L_2, распознаваемый конечным автоматом $M = \langle Q, \Sigma, \Delta, I, F\rangle$, где $Q = \{1, 2, 3\}$, $I = \{1\}$, $F = \{3\}$,

$$\Delta = \{\langle 1, a, 1\rangle,\ \langle 1, b, 1\rangle,\ \langle 1, a, 2\rangle,\ \langle 2, b, 3\rangle,\ \langle 3, a, 3\rangle,\ \langle 3, b, 3\rangle\}.$$



Тогда язык $L_1 \cap L_2$ порождается контекстно-свободной грамматикой $\overline{S} \to S_{13}$, $S_{11} \to aaS_{11}aa$, $S_{12} \to aaS_{11}aa$, $S_{13} \to aaS_{13}aa$, $S_{13} \to aaS_{23}aa$, $S_{33} \to aaS_{33}aa$, $S_{11} \to bS_{11}b$, $S_{13} \to bS_{13}b$, $S_{13} \to bS_{12}b$, $S_{23} \to bS_{33}b$, $S_{33} \to bS_{33}b$, $S_{11} \to \varepsilon$, $S_{33} \to \varepsilon$. Эту грамматику можно упростить, заменив S_{11} и S_{33} на один символ.

Задачи

4.6.1. Найти контекстно-свободную грамматику для языка $L_1 \cap L_2$, где L_1 порождается грамматикой $F \to aFa$, $F \to bFb$, $F \to \varepsilon$, а язык L_2 порождается грамматикой $K \to aM$, $M \to aK$, $K \to bM$, $K \to bK$, $M \to \varepsilon$.

4.6.2. Найти контекстно-свободную грамматику для языка $L_1 \cap L_2$, где L_1 порождается грамматикой $F \to aFa$, $F \to bFb$, $F \to \varepsilon$, а язык L_2 порождается грамматикой $K \to aM$, $M \to aK$, $K \to bK$, $M \to bK$, $M \to \varepsilon$.

4.6.3. Найти контекстно-свободную грамматику для языка $L_1 \cap L_2$, где L_1 порождается грамматикой $F \to aFa$, $F \to bFb$, $F \to \varepsilon$, а язык L_2 порождается грамматикой $K \to aM$, $M \to aK$, $K \to bM$, $K \to bK$, $K \to \varepsilon$.

4.6.4. Найти контекстно-свободную грамматику для языка $L_1 \cap L_2$, где язык L_1 порождается грамматикой $F \to a$, $F \to bF$, $F \to cFF$, а язык L_2 порождается грамматикой $K \to aM$, $K \to cM$, $M \to aK$, $M \to bK$, $M \to \varepsilon$.

4.6.5. Найти контекстно-свободную грамматику для языка $L_1 \cap L_2$, где язык L_1 порождается грамматикой $F \to aFb$, $F \to aFa$, $F \to \varepsilon$, а язык L_2 порождается грамматикой $K_1 \to aK_2$, $K_2 \to aK_1$, $K_2 \to bK_3$, $K_3 \to aK_1$, $K_3 \to \varepsilon$.

4.6.6. Найти контекстно-свободную грамматику для языка $L_1 \cap L_2$, где язык L_1 порождается грамматикой $F \to a$, $F \to abF$, $F \to baFF$, а язык L_2 порождается грамматикой $K \to baaK$, $K \to baaaK$, $K \to \varepsilon$.

4.6.7. Найти контекстно-свободную грамматику для языка $L_1 - L_2$, где язык L_1 порождается грамматикой $F \to a$, $F \to GF$, $G \to b$, $G \to cF$, а язык L_2 порождается грамматикой $S \to STT$, $S \to UbT$, $T \to a$, $T \to b$, $T \to c$, $U \to UT$, $U \to \varepsilon$.

4.6.8. Эквивалентны ли грамматика $J \to F_{12}$, $J \to F_{22}$, $F_{11} \to cF_{21}F_{11}$, $F_{11} \to G_{12}F_{21}$, $F_{12} \to a$, $F_{12} \to cF_{21}F_{12}$,

$F_{12} \to G_{12}F_{22}$, $F_{21} \to a$, $F_{21} \to cF_{11}F_{11}$, $F_{21} \to cF_{12}F_{21}$, $F_{22} \to cF_{11}F_{12}$, $F_{22} \to cF_{12}F_{22}$, $G_{12} \to b$, $G_{12} \to cF_{22}$ и грамматика $S \to O$, $S \to E$, $O \to a$, $O \to bE$, $O \to cEE$, $O \to caO$, $O \to ccOaO$, $E \to cOE$, $E \to ccEaO$, $E \to cbcOaO$, $E \to cbaO$?

4.6.9. Эквивалентны ли грамматика $J \to F_{12}$, $J \to F_{22}$, $F_{11} \to cF_{21}F_{11}$, $F_{11} \to G_{12}F_{21}$, $F_{12} \to a$, $F_{12} \to cF_{21}F_{12}$, $F_{12} \to G_{12}F_{22}$, $F_{21} \to a$, $F_{21} \to cF_{11}F_{11}$, $F_{21} \to cF_{12}F_{21}$, $F_{22} \to cF_{11}G_{12}F_{22}$, $F_{22} \to cF_{12}F_{22}$, $G_{12} \to b$, $G_{12} \to cF_{22}$ и грамматика $S \to O$, $S \to E$, $O \to a$, $O \to bE$, $O \to cEE$, $O \to caO$, $O \to ccOaO$, $E \to cOE$, $E \to ccEaO$, $E \to cbcOabE$, $E \to cbaO$?

4.6.10. Является ли контекстно-свободным язык $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\} - \{(ab)^n \mid n > 0\}$?

4.6.11. Является ли контекстно-свободным язык $\{a^i b^{3j} a b^k a^l b^{2l} a^j b^{k+1} a^i \mid i \geqslant 0,\ j \geqslant 0,\ k \geqslant 0,\ l \geqslant 0\}$?

4.6.12. Является ли контекстно-свободным язык $\{a^n b^n a^k b^n \mid n \geqslant 0,\ k \geqslant 0\}$?

4.6.13. Является ли контекстно-свободным язык $\{a^k b^l c^m a^m b^k c^l \mid k \geqslant 0,\ l \geqslant 0,\ m \geqslant 0\}$?

4.6.14. Является ли контекстно-свободным язык $\{(ab)^{n^2} \mid n \geqslant 0\} \cup \{aaw \mid w \in \{a,b\}^*\}$?

4.6.15. Является ли контекстно-свободным язык $\{w \in \{a,b,c\}^* \mid 2|w|_a = |w|_b + |w|_c,\ |w|_b \leqslant |w|_a\}$?

4.6.16. Является ли контекстно-свободным язык $\{w \in \{a,b,c\}^* \mid 2|w|_a = |w|_b + |w|_c,\ |w|_b \leqslant 3 \leqslant |w|_a\}$?

4.6.17. Является ли контекстно-свободным язык $\{a^n b^n a^n \mid n \geqslant 0\} \cup \{a^k b v_1 a b v_2 \mid k \geqslant 0,\ v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*\}$?

4.6.18. Является ли контекстно-свободным язык $\{v_1 v_2 v_1 v_3 \mid v_1 \in \{a,b\}^*,\ v_2 \in \{a,b\}^*,\ v_3 \in \{a,b\}^*,\ v_2 = v_2^{\mathrm{R}},\ v_3 = v_3^{\mathrm{R}}\}$?

4.6.19. Является ли непустым пересечение языка, порождаемого грамматикой $S \to a$, $S \to abababab S$, и языка, порождаемого грамматикой $M \to aMaTRa$, $M \to R$, $R \to bRaTb$, $R \to T$, $T \to aMb$, $T \to ba$, $T \to ab$?

4.6.20. Является ли бесконечным пересечение языка, порождаемого грамматикой $S \to a$, $S \to abS$, и языка, порождаемого грамматикой $M \to aMaRa$, $M \to R$, $R \to bRaTb$, $R \to T$, $T \to aMb$, $T \to ba$, $T \to ab$?

4.6.21. Существуют ли такие языки L_1 и L_2, что $L_1 \subseteq L_2$, L_2 является автоматным и L_1 не является контекстно-свободным?

4.6.22. Существуют ли такие языки L_1 и L_2, что L_1 является линейным, L_2 является автоматным и $L_1 \cup L_2$ не является линейным?

4.6.23. Существуют ли такие языки L_1 и L_2, что L_1 является линейным, L_2 является автоматным и $L_1 \cap L_2$ не является линейным?

4.6.24. Существуют ли такие языки L_1 и L_2, что L_1 является линейным, L_2 является линейным и $L_1 \cap L_2$ не является контекстно-свободным?

4.6.25. Существует ли над алфавитом $\{a, b\}$ такой линейный язык L, что язык $\{ubv \mid u \in L,\ uv \in L\}$ не является контекстно-свободным?

4.6.26. Существуют ли праволинейная грамматика G_1 и однозначная контекстно-свободная грамматика G_2, такие что язык $L(G_1) \cap L(G_2)$ — существенно неоднозначный контекстно-свободный язык?

4.7. Теорема Парика

[20, 9.7], [6, 5.1, 5.2], [7, 4.3], [30, 3.5.2], [2, с. 239]

Замечание 1. В этом разделе предполагается, что зафиксирован некоторый линейный порядок на алфавите Σ. Пусть $\Sigma = \{a_1, \ldots, a_n\}$.

Определение 2. Через Ψ_Σ будем обозначать функцию из Σ^* в $\mathbb{N}^n$, определённую следующим образом: $\Psi_\Sigma(w) \risingdotseq \langle |w|_{a_1}, \ldots, |w|_{a_n} \rangle$. Аналогично, каждому языку $L \subseteq \Sigma^*$ ставится в соответствие множество $\Psi_\Sigma(L) \subseteq \mathbb{N}^n$, определённое следующим образом:

$$\Psi_\Sigma(L) \risingdotseq \{\Psi_\Sigma(w) \mid w \in L\}.$$

Пример 3. Пусть $\Sigma = \{a_1, a_2\}$ и $L = \{a_1, a_1a_2a_2, a_2a_2a_1\}$. Тогда $\Psi_\Sigma(L) = \{\langle 1, 0\rangle, \langle 1, 2\rangle\}$.

Определение 4. Пусть $B \subseteq \mathbb{N}^n$ и $P \subseteq \mathbb{N}^n$. Тогда через $\mathrm{L}(B,P)$ обозначается множество

$$\{b + p_1 + \ldots + p_k \mid b \in B,\ k \geqslant 0,\ p_1, \ldots, p_k \in P\}.$$

При этом множество B называется *системой предпериодов* множества $\mathrm{L}(B,P)$. Множество P называется *системой периодов* множества $\mathrm{L}(B,P)$.

Определение 5. Множество $A \subseteq \mathbb{N}^n$ называется *линейным* (linear), если $A = \mathrm{L}(B,P)$ для некоторых конечных множеств B и P.

Определение 6. Множество $A \subseteq \mathbb{N}^n$ называется *полулинейным* (semilinear), если оно является объединением конечного числа линейных множеств.

Теорема 7 (Теорема Пари́ка)**.** *Если язык $L \subseteq \Sigma^*$ является контекстно-свободным, то множество $\Psi_\Sigma(L)$ является полулинейным.*

Доказательство можно найти в [6, с. 207—211].

Пример 8. Пусть $\Sigma = \{a, b\}$. Рассмотрим язык $L = \{a^m b^n \mid m > n$ или m простое$\}$. Можно проверить, что множество $\Psi_\Sigma(L)$ не является полулинейным. Следовательно, язык L не является контекстно-свободным.

Теорема 9. *Если множество $A \subseteq \mathbb{N}^n$ является полулинейным, то существует такой автоматный язык L, что $A = \Psi_\Sigma(L)$.*

Задачи

4.7.1. Является ли контекстно-свободным язык $\{a^{m^2} b^n \mid m \geqslant 0,\ n \geqslant 0\} \cup \{a^m b^n \mid 0 \leqslant n < m\}$?

4.7.2. Является ли контекстно-свободным язык $\{w \in \{a,b\}^* \mid |w|_a^2 + |w|_b^2 < 50\}$?

4.7.3. Является ли контекстно-свободным язык $\{w \in \{a,b\}^* \mid |w|_a^2 + |w|_b^2 > 9\}$?

4.7.4. Является ли контекстно-свободным язык $\{w \in \{a,b\}^* \mid |w|_a^2 - |w|_b^2 > 9\}$?

4.7.5. Является ли контекстно-свободным язык $\{w \in \{a,b\}^* \mid |w|_a^2 - |w|_b > 9\}$?

4.7.6. Является ли контекстно-свободным язык $\{w \in \{a,b\}^* \mid |w|_a - |w|_b^2 > 9\}$?

5. Автоматы с магазинной памятью

5.1. Определение автомата с магазинной памятью

[20, 10.1], [6, 2.4], [2, 2.5.1, 2.5.2], [27, 6.1, 6.2], [7, 4.5], [8, с. 136—149], [9, с. 418—420], [24, 4.1], [18, I.5.1, I.5.2], [19, 6.1], [25, с. 36—38], [16, 5.2.5, 5.2.6], [4, с. 109—116], [11, 9.1, 9.2], [31, 3.3], [32, 2.2], [22, с. 79—83], [13, 3.2], [17, 5.1, 5.3, 5.4]

Определение 1. *Автомат с магазинной памятью* (*МП-автомат, магазинный автомат, стековый автомат,* pushdown automaton) — это шестёрка $M = \langle Q, \Sigma, \Gamma, \Delta, I, F\rangle$, где Q, Σ, Γ и Δ — конечные множества, $I \subseteq Q$, $F \subseteq Q$ и

$$\Delta \subseteq (Q \times \Sigma^* \times \Gamma^*) \times (Q \times \Gamma^*).$$

Здесь Q — *множество состояний*, Σ — *входной алфавит*, Γ — *алфавит магазинной памяти* (stack alphabet), Δ — *множество переходов* (transition relation), элементы I называются *начальными* состояниями, элементы F — *заключительными* или *допускающими* состояниями.

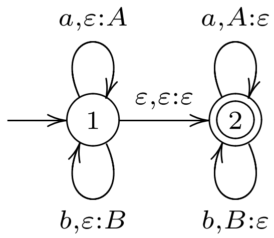
Пример 2. Пусть $Q = \{1, 2\}$, $\Sigma = \{a, b\}$, $\Gamma = \{A, B\}$,

$$\Delta = \{\langle\langle 1, a, \varepsilon\rangle, \langle 1, A\rangle\rangle,\ \langle\langle 1, b, \varepsilon\rangle, \langle 1, B\rangle\rangle,\ \langle\langle 1, \varepsilon, \varepsilon\rangle, \langle 2, \varepsilon\rangle\rangle,$$
$$\langle\langle 2, a, A\rangle, \langle 2, \varepsilon\rangle\rangle,\ \langle\langle 2, b, B\rangle, \langle 2, \varepsilon\rangle\rangle\},$$

$I = \{1\}$, $F = \{2\}$. Тогда $\langle Q, \Sigma, \Gamma, \Delta, I, F\rangle$ — МП-автомат.

Замечание 3. МП-автоматы можно изображать в виде *диаграмм состояний*. На диаграмме каждое состояние обозначается кружком, а переход — стрелкой. Каждое начальное состояние распознаётся по ведущей в него короткой стрелке. Каждое допускающее состояние отмечается на диаграмме двойным кружком. Стрелка с пометкой $x, \beta : \gamma$, ведущая из p в q, показывает, что $\langle\langle p, x, \beta\rangle, \langle q, \gamma\rangle\rangle$ является переходом данного МП-автомата.

Пример 4. Ниже приведена диаграмма МП-автомата из примера 2.



Определение 5. *Конфигурацией* МП-автомата называется любая тройка $\langle q, w, \alpha\rangle$, где $q \in Q$, $w \in \Sigma^*$, $\alpha \in \Gamma^*$.

Определение 6. Определим на множестве всех конфигураций МП-автомата $M = \langle Q, \Sigma, \Gamma, \Delta, I, F\rangle$ бинарное отношение $\underset{M}{\vdash}$ (*такт работы*) следующим образом. Если $\langle\langle p, x, \beta\rangle, \langle q, \gamma\rangle\rangle \in \Delta$, $w \in \Sigma^*$ и $\alpha \in \Gamma^*$, то $\langle p, xw, \beta\alpha\rangle \underset{M}{\vdash} \langle q, w, \gamma\alpha\rangle$.

Замечание 7. Обычно из контекста ясно, о каком МП-автомате идёт речь. Тогда вместо $\underset{M}{\vdash}$ будем писать $\vdash$.

Определение 8. Бинарное отношение $\overset{*}{\vdash}$ определяется как рефлексивное, транзитивное замыкание отношения $\vdash$.

Пример 9. Для МП-автомата из примера 2 выполняется $\langle 1, abba, \varepsilon\rangle \overset{*}{\vdash} \langle 1, ba, BA\rangle$ и $\langle 1, abba, \varepsilon\rangle \overset{*}{\vdash} \langle 2, \varepsilon, \varepsilon\rangle$.

Определение 10. Слово $w \in \Sigma^*$ *допускается* МП-автоматом $M = \langle Q, \Sigma, \Gamma, \Delta, I, F\rangle$, если найдутся такие состояния $s \in I$ и $q \in F$, что $\langle s, w, \varepsilon\rangle \overset{*}{\vdash} \langle q, \varepsilon, \varepsilon\rangle$.

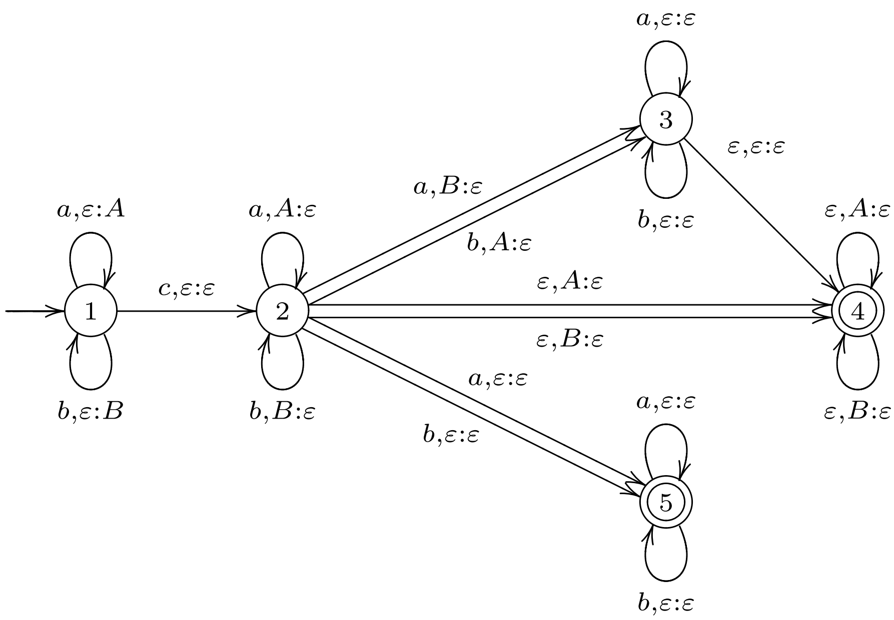
Определение 11. Язык, *распознаваемый* МП-автоматом, — множество всех слов, допускаемых этим МП-автоматом. Язык, распознаваемый МП-автоматом M, обозначается $L(M)$.

Пример 12. Пусть M — МП-автомат из примера 2. Тогда $L(M) = \{uu^{\mathrm{R}} \mid u \in \{a, b\}^*\}$.

Пример 13. Пусть $\Sigma = \{a, b, c\}$. Рассмотрим МП-автомат $M = \langle Q, \Sigma, \Gamma, \Delta, I, F\rangle$, где $Q = \{1, 2, 3, 4, 5\}$, $\Gamma = \{A, B\}$, $I = \{1\}$, $F = \{4, 5\}$ и

$$\begin{aligned}\Delta = \{&\langle\langle 1, a, \varepsilon\rangle, \langle 1, A\rangle\rangle,\ \langle\langle 1, b, \varepsilon\rangle, \langle 1, B\rangle\rangle,\\ &\langle\langle 1, c, \varepsilon\rangle, \langle 2, \varepsilon\rangle\rangle,\\ &\langle\langle 2, a, A\rangle, \langle 2, \varepsilon\rangle\rangle,\ \langle\langle 2, b, B\rangle, \langle 2, \varepsilon\rangle\rangle,\\ &\langle\langle 2, a, B\rangle, \langle 3, \varepsilon\rangle\rangle,\ \langle\langle 2, b, A\rangle, \langle 3, \varepsilon\rangle\rangle,\end{aligned}$$

$$\begin{aligned}
&\langle\langle 2,\varepsilon,A\rangle,\langle 4,\varepsilon\rangle\rangle,\ \langle\langle 2,\varepsilon,B\rangle,\langle 4,\varepsilon\rangle\rangle,\\
&\langle\langle 2,b,\varepsilon\rangle,\langle 5,\varepsilon\rangle\rangle,\ \langle\langle 2,a,\varepsilon\rangle,\langle 5,\varepsilon\rangle\rangle,\\
&\langle\langle 3,a,\varepsilon\rangle,\langle 3,\varepsilon\rangle\rangle,\ \langle\langle 3,b,\varepsilon\rangle,\langle 3,\varepsilon\rangle\rangle,\\
&\langle\langle 3,\varepsilon,\varepsilon\rangle,\langle 4,\varepsilon\rangle\rangle,\\
&\langle\langle 4,\varepsilon,A\rangle,\langle 4,\varepsilon\rangle\rangle,\ \langle\langle 4,\varepsilon,B\rangle,\langle 4,\varepsilon\rangle\rangle,\\
&\langle\langle 5,a,\varepsilon\rangle,\langle 5,\varepsilon\rangle\rangle,\ \langle\langle 5,b,\varepsilon\rangle,\langle 5,\varepsilon\rangle\rangle\}.
\end{aligned}$$



Тогда $L(M) = \{ucv \mid u \in \{a,b\}^*,\ v \in \{a,b\}^*,\ u \neq v^{\mathrm{R}}\}$.

Определение 14. Два МП-автомата *эквивалентны*, если они распознают один и тот же язык.

Задачи

5.1.1. Найти автомат с магазинной памятью, распознающий язык $\{a^n b^n \mid n \geqslant 0\}$.

5.1.2. Найти автомат с магазинной памятью, распознающий язык $\{a^i b^j c^k \mid i = j$ или $j = k\}$.

5.1.3. Найти автомат с магазинной памятью, распознающий язык $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$.

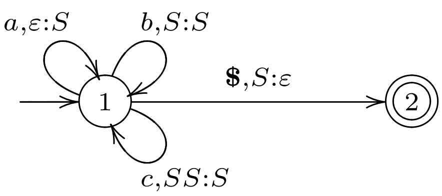
5.1.4. Найти автомат с магазинной памятью, распознающий язык $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c = 1\}$.

5.1.5. Найти автомат с магазинной памятью, распознающий язык $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c \leqslant 2\}$.

5.1.6. Найти автомат с магазинной памятью, распознающий язык $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c \geqslant 2\}$.

5.1.7. Найти автомат с магазинной памятью, распознающий язык $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c \,\vdots\, 2\}$.

5.1.8. Описать язык над алфавитом $\{a,b,c,\$\}$, распознаваемый следующим МП-автоматом.



5.2. Характеризация контекстно-свободных языков

[20, 10.2], [32, 2.2], [6, 2.5], [2, 2.5.3], [27, 6.3], [7, 4.5], [8, с. 149—150], [9, с. 420—422], [24, 4.1], [18, I.5.3], [4, с. 116—120], [11, 9.3, 15.3], [31, 3.4], [22, с. 84—87], [19, 6.2, 6.3], [13, 3.2]

Теорема 1. *Если язык L является контекстно-свободным, то существует МП-автомат, распознающий этот язык.*

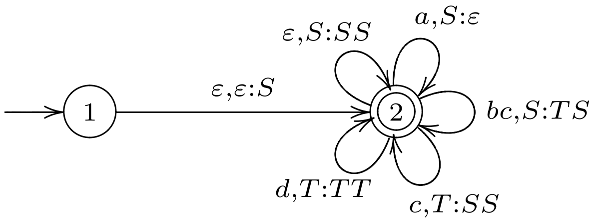
Доказательство. Пусть язык L порождается контекстно-свободной грамматикой $\langle N, \Sigma, P, S\rangle$, в которой каждое правило имеет вид $A \to w\beta$, где $A \in N$, $w \in \Sigma^*$ и $\beta \in N^*$ (в силу теоремы 3 на с. 35 такая грамматика существует). Положим $\Gamma = N$, $Q = \{1,2\}$, $I = \{1\}$, $F = \{2\}$ и

$$\Delta = \{\langle\langle 1,\varepsilon,\varepsilon\rangle, \langle 2,S\rangle\rangle\} \cup \{\langle\langle 2,w,A\rangle, \langle 2,\beta\rangle\rangle \mid (A \to w\beta) \in P\}.$$

Можно доказать, что $\langle 2,u,S\rangle \overset{*}{\vdash} \langle 2,\varepsilon,\alpha\rangle$ тогда и только тогда, когда существует левосторонний вывод $S \underset{\mathrm{lm}}{\overset{*}{\Rightarrow}} u\alpha$ (здесь $u \in \Sigma^*$ и $\alpha \in N^*$). □

Пример 2. Пусть $\Sigma = \{a,b,c,d\}$. Контекстно-свободная грамматика $S \to SS$, $S \to a$, $S \to bcTS$, $T \to cSS$, $T \to dTT$ и МП-автомат $\langle\{1,2\}, \Sigma, \{S,T\}, \Delta, \{1\}, \{2\}\rangle$, где

$$\Delta = \{\langle\langle 1,\varepsilon,\varepsilon\rangle, \langle 2,S\rangle\rangle,\ \langle\langle 2,\varepsilon,S\rangle, \langle 2,SS\rangle\rangle,\ \langle\langle 2,a,S\rangle, \langle 2,\varepsilon\rangle\rangle,$$
$$\langle\langle 2,bc,S\rangle, \langle 2,TS\rangle\rangle,\ \langle\langle 2,c,T\rangle, \langle 2,SS\rangle\rangle,\ \langle\langle 2,d,T\rangle, \langle 2,TT\rangle\rangle\},$$



задают один и тот же язык.

Лемма 3. *Каждый МП-автомат эквивалентен некоторому МП-автомату $\langle Q, \Sigma, \Gamma, \Delta, I, F\rangle$, где $|I| = 1$, $|F| = 1$ и каждый переход $\langle\langle p, x, \beta\rangle, \langle q, \gamma\rangle\rangle \in \Delta$ удовлетворяет требованию $|\beta| + |\gamma| = 1$.*

Теорема 4. *Если язык L распознаётся некоторым МП-автоматом, то L является контекстно-свободным.*

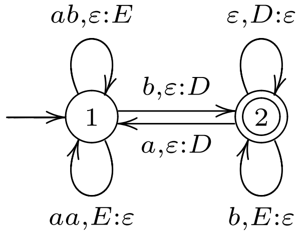
Доказательство. Пусть язык L распознаётся МП-автоматом $\langle Q, \Sigma, \Gamma, \Delta, I, F\rangle$. Без ограничения общности можно считать, что $I = \{q_{\mathrm{s}}\}$, $F = \{q_{\mathrm{a}}\}$ и каждый переход $\langle\langle p, x, \beta\rangle, \langle q, \gamma\rangle\rangle \in \Delta$ удовлетворяет требованию $|\beta| + |\gamma| = 1$. Построим искомую контекстно-свободную грамматику $\langle N, \Sigma, P, S\rangle$, положив $N = \{A_{p,q} \mid p \in Q,\ q \in Q\}$, $S = A_{q_{\mathrm{s}},q_{\mathrm{a}}}$ и

$$\begin{aligned}P = \{A_{p,p} \to \varepsilon \mid p \in Q\} \cup \\ \cup \{A_{p,t} \to xA_{q,r}yA_{s,t} \mid \langle\langle p, x, \varepsilon\rangle, \langle q, C\rangle\rangle \in \Delta, \\ \langle\langle r, y, C\rangle, \langle s, \varepsilon\rangle\rangle \in \Delta,\ C \in \Gamma,\ t \in Q\}.\end{aligned}$$

Можно доказать, что $\langle p, x, \varepsilon\rangle \overset{*}{\vdash} \langle q, \varepsilon, \varepsilon\rangle$ тогда и только тогда, когда $A_{p,q} \overset{*}{\Rightarrow} x$ (здесь $x \in \Sigma^*$). $\square$

Пример 5. МП-автомат $M = \langle\{1, 2\}, \Sigma, \Gamma, \Delta, \{1\}, \{2\}\rangle$, где $\Sigma = \{a, b\}$, $\Gamma = \{D, E\}$,

$$\begin{aligned}\Delta = \{&\langle\langle 1, ab, \varepsilon\rangle, \langle 1, E\rangle\rangle,\ \langle\langle 1, aa, E\rangle, \langle 1, \varepsilon\rangle\rangle,\ \langle\langle 1, b, \varepsilon\rangle, \langle 2, D\rangle\rangle, \\ &\langle\langle 2, a, \varepsilon\rangle, \langle 1, D\rangle\rangle,\ \langle\langle 2, \varepsilon, D\rangle, \langle 2, \varepsilon\rangle\rangle,\ \langle\langle 2, b, E\rangle, \langle 2, \varepsilon\rangle\rangle\},\end{aligned}$$



и контекстно-свободная грамматика $S \to abTaaS$, $S \to abSbU$, $S \to bUU$, $T \to abTaaT$, $T \to \varepsilon$, $U \to aSU$, $U \to \varepsilon$ задают один и тот же язык. Здесь S, T и U соответствуют символам $A_{1,2}$, $A_{1,1}$ и $A_{2,2}$ из доказательства теоремы 4.

Задачи

5.2.1. Найти автомат с магазинной памятью, распознающий язык, порождаемый грамматикой $F \to aFb$, $F \to bFa$, $F \to c$.

5.2.2. Найти автомат с магазинной памятью, распознающий язык, порождаемый грамматикой $F \to aFb$, $F \to aFa$, $F \to a$.

В задачах 5.2.3—5.2.24 надо найти контекстно-свободную грамматику, порождающую данный язык.

5.2.3. $\{ua^nb^n \mid u \in \{a,b\}^+,\ u = u^{\mathrm{R}},\ n \geqslant 1\}$.

5.2.4. $\{v_1v_1^{\mathrm{R}}a^nv_2 \mid v_1 \in \{a,b\}^+,\ v_2 \in \{a,b\}^+,\ v_2 = v_2^{\mathrm{R}},\ n \geqslant 0\}$.

5.2.5. $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$.

5.2.6. $\{w \in \{a,b\}^+ \mid |w|_a = |w|_b\}$.

5.2.7. $\{w \in \{a,b\}^* \mid |w|_a \geqslant |w|_b\}$.

5.2.8. $\{w \in \{a,b\}^+ \mid |w|_a \geqslant |w|_b\}$.

5.2.9. $\{w \in \{a,b\}^* \mid |w|_a > |w|_b\}$.

5.2.10. $\{w \in \{a,b\}^* \mid |w|_a = 2|w|_b\}$.

5.2.11. $\{w \in \{a,b\}^+ \mid |w|_a = 2|w|_b\}$.

5.2.12. $\{w \in \{a,b\}^* \mid |w|_a = 2|w|_b\} - (b(a+b)^* + (a+b)^*bb(a+b)^*)$.

5.2.13. $\{w \in \{a,b,c,d\}^* \mid |w|_a + |w|_b = |w|_c + |w|_d\}$.

5.2.14. $\{w \in \{a,b\}^* \mid |w|_a \neq |w|_b\}$.

5.2.15. $\{w \in \{a,b\}^* \mid 2|w|_a = 3|w|_b\}$.

5.2.16. $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c = 1\}$.

5.2.17. $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c = 2\}$.

5.2.18. $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c \leqslant 2\}$.

5.2.19. $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c \geqslant 2\}$.

5.2.20. $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b,\ |w|_c \,\vdots\, 2\}$.

5.2.21. $\{w \in \{a,b,c\}^* \mid |w|_a \geqslant |w|_b,\ |w|_c \leqslant 2\}$.

5.2.22. $\{w \in \{a,b,c\}^* \mid |w|_a \geqslant |w|_b,\ |w|_c \geqslant 2\}$.

5.2.23. $\{w \in \{a,b,c\}^* \mid |w|_a \geqslant |w|_b,\ |w|_c \,\vdots\, 2\}$.

5.2.24. $\{w \in \{a,b\}^* \mid f(w) = 0\}$, где $f(\varepsilon) \leftrightharpoons 0$, $f(xa) \leftrightharpoons 1 - f(x)$ и $f(xb) \leftrightharpoons -1 - f(x)$ для любого $x \in \{a,b\}^*$.

6. Дополнительные свойства контекстно-свободных языков

6.1. Деление контекстно-свободных языков

[20, 11.1], [18, I.9.2], [2, с. 236], [31, с. 148—149]

Теорема 1. *Пусть L_1 — контекстно-свободный язык над алфавитом Σ и L_2 — автоматный язык над алфавитом Σ. Тогда язык $\{x \in \Sigma^* \mid (\exists y \in L_2)\, xy \in L_1\}$ является контекстно-свободным.*

Доказательство можно найти в [20, 11.1].

Пример 2. Пусть $\Sigma = \{a, b\}$, $L_1 = \{w \in \Sigma^* \mid w = w^{\mathrm{R}}\}$ и $L_2 = \{v_1 aabb v_2 \mid v_1 \in \Sigma^*,\ v_2 \in \Sigma^*\}$. Тогда

$$\{x \in \Sigma^* \mid (\exists y \in L_2)\, xy \in L_1\} = \Sigma^*.$$

Замечание 3. Пусть $\Sigma = \{a, b, c\}$ и $L \subseteq \Sigma^*$. Язык $L \cdot \{c\}$ является контекстно-свободным тогда и только тогда, когда язык L является контекстно-свободным.

Задачи

6.1.1. Найти контекстно-свободную грамматику для языка $\{x \in \{a, b\}^* \mid (\exists y \in L_2)\, xy \in L_1\}$, где $L_1 = \{w \in \{a, b\}^* \mid |w|_a = |w|_b\}$ и $L_2 = \{a^k (ab)^m \mid k \geqslant 0,\ m \geqslant 1\}$.

6.1.2. Найти контекстно-свободную грамматику для языка $\{x \in \{a, b\}^* \mid (\exists y \in L_2)\, xy \in L_1\}$, где $L_1 = \{a^k b^m \mid k \geqslant 0,\ m \geqslant 1\}$ и $L_2 = \{w \in \{a, b\}^* \mid |w|_a = |w|_b\}$.

6.1.3. Найти контекстно-свободную грамматику для языка $\{y \in \{a, b\}^* \mid (\exists x \in L_2)\, xy \in L_1\}$, где $L_1 = \{a^k b^m \mid k \geqslant 0,\ m \geqslant 1\}$ и $L_2 = \{w \in \{a, b\}^* \mid |w|_a > |w|_b\}$.

6.1.4. Найти контекстно-свободную грамматику для языка $\{y \in \{a, b\}^* \mid (\exists x \in L_2)\, xy \in L_1\}$, где $L_1 = \{a^m b^m \mid m \geqslant 1\}^*$ и $L_2 = \{w \in \{a, b\}^* \mid |w|_a < |w|_b\}$.

6.1.5. Существует ли такой контекстно-свободный язык $L \subseteq \Sigma^*$, что язык $\mathrm{Subw}(L)$ не является контекстно-свободным?

6.1.6. Существует ли такой контекстно-свободный язык L над алфавитом $\{a, b\}$, что язык $\{w \mid (\exists n \geqslant 0)\, wa^{2n} \in L\}$ не является контекстно-свободным?

6.1.7. Существует ли такой контекстно-свободный язык L над алфавитом $\{a, b\}$, что язык

$$\{w \in \{a, b\}^* \mid (\exists m \geqslant 0)\, (\exists n \geqslant 0)\, (ab)^m wb^{3n} \in L\}$$

не является контекстно-свободным?

6.1.8. Существует ли такой контекстно-свободный язык L над алфавитом $\{a, b, c\}$, что язык $\{w \mid waba \in L\}$ не является контекстно-свободным?

6.1.9. Существует ли такой язык L над алфавитом $\{a, b, c\}$, что язык $\{w \mid waba \in L\}$ является контекстно-свободным, но язык L не является контекстно-свободным?

6.2. Гомоморфизмы и контекстно-свободные языки

[20, 11.2], [2, 2.6.2], [27, 7.3.5], [7, 4.2], [18, I.9.2], [31, с. 148]

Теорема 1. *Для любого гомоморфизма $h\colon \Sigma_1^* \to \Sigma_2^*$ и контекстно-свободного языка $L \subseteq \Sigma_1^*$ язык $h(L)$ является контекстно-свободным.*

Пример 2. Пусть $\Sigma_1 = \{a, b, c\}$ и $\Sigma_2 = \{a, b\}$. Рассмотрим контекстно-свободный язык L, порождаемый грамматикой $S \to aSSc$, $S \to b$, и гомоморфизм $h\colon \Sigma_1^* \to \Sigma_2^*$, заданный равенствами $h(a) = a$, $h(b) = bba$ и $h(c) = a$. Тогда язык $h(L)$ порождается контекстно-свободной грамматикой $S \to aSSa$, $S \to bba$.

Теорема 3. *Для любого гомоморфизма $h\colon \Sigma_1^* \to \Sigma_2^*$ и контекстно-свободного языка $L \subseteq \Sigma_2^*$ язык $h^{-1}(L)$ является контекстно-свободным.*

Доказательство. Введём обозначение

$$\mathcal{A} = \{x \in \Sigma_2^* \mid x \sqsupset h(a) \text{ для некоторого } a \in \Sigma_1\}.$$

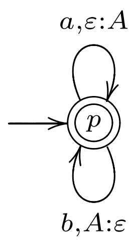
Пусть язык L распознаётся МП-автоматом $\langle Q, \Sigma_2, \Gamma, \Delta, I, F\rangle$. Без ограничения общности можно считать, что для каждого перехода $\langle\langle p,x,\beta\rangle,\langle q,\gamma\rangle\rangle \in \Delta$ выполняется неравенство $|x| \leqslant 1$. Можно проверить, что в этом случае язык $h^{-1}(L)$ распознаётся МП-автоматом $\langle Q',\Sigma_1,\Gamma,\Delta',I',F'\rangle$, где

$$\begin{aligned}
Q' &= \mathcal{A} \times Q,\\
I' &= \{\varepsilon\} \times I,\\
F' &= \{\varepsilon\} \times F,\\
\Delta' &= \{\langle\langle\langle\varepsilon,p\rangle,a,\varepsilon\rangle,\langle\langle h(a),p\rangle,\varepsilon\rangle\rangle \mid a \in \Sigma_1,\ p \in Q\} \cup\\
&\quad \cup \{\langle\langle\langle xw,p\rangle,\varepsilon,\beta\rangle,\langle\langle w,q\rangle,\gamma\rangle\rangle \mid \langle\langle p,x,\beta\rangle,\langle q,\gamma\rangle\rangle \in \Delta,\ xw \in \mathcal{A}\}.
\end{aligned}$$

$\square$

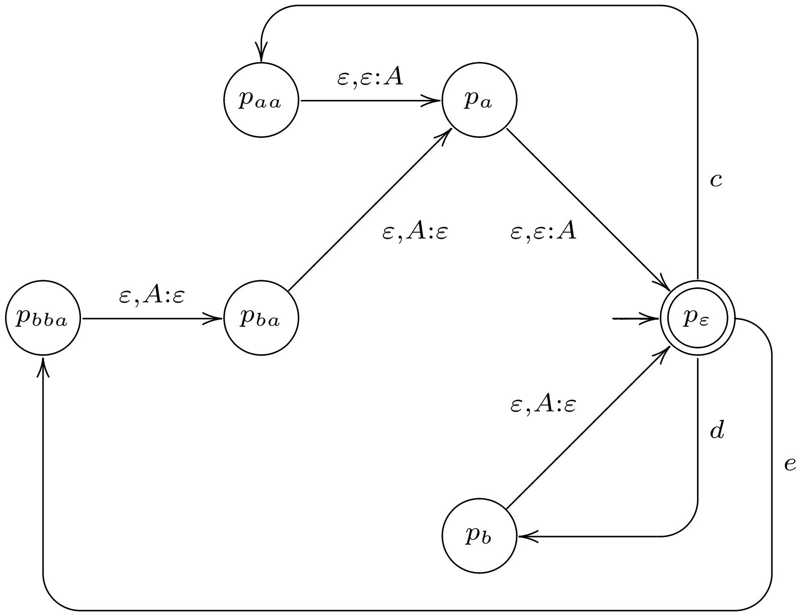
Пример 4. Пусть $\Sigma_1 = \{c,d,e\}$ и $\Sigma_2 = \{a,b\}$. Рассмотрим гомоморфизм $h\colon \Sigma_1^* \to \Sigma_2^*$, заданный равенствами $h(c) = aa$, $h(d) = b$ и $h(e) = bba$. Пусть контекстно-свободный язык L распознаётся МП-автоматом $\langle Q,\Sigma_2,\Gamma,\Delta,I,F\rangle$, где $Q = \{p\}$, $\Gamma = \{A\}$, $I = \{p\}$, $F = \{p\}$,

$$\Delta = \{\langle\langle p,a,\varepsilon\rangle,\langle p,A\rangle\rangle,\ \langle\langle p,b,A\rangle,\langle p,\varepsilon\rangle\rangle\}.$$



Тогда язык $h^{-1}(L)$ распознаётся МП-автоматом $\langle Q',\Sigma_1,\Gamma,\Delta',I',F'\rangle$, где $Q' = \{p_\varepsilon,p_a,p_{aa},p_b,p_{ba},p_{bba}\}$, $I' = \{p_\varepsilon\}$, $F' = \{p_\varepsilon\}$ и

$$\begin{aligned}
\Delta' = \{&\langle\langle p_\varepsilon,c,\varepsilon\rangle,\langle p_{aa},\varepsilon\rangle\rangle,\ \langle\langle p_\varepsilon,d,\varepsilon\rangle,\langle p_b,\varepsilon\rangle\rangle,\ \langle\langle p_\varepsilon,e,\varepsilon\rangle,\langle p_{bba},\varepsilon\rangle\rangle,\\
&\langle\langle p_a,\varepsilon,\varepsilon\rangle,\langle p_\varepsilon,A\rangle\rangle,\ \langle\langle p_{aa},\varepsilon,\varepsilon\rangle,\langle p_a,A\rangle\rangle,\ \langle\langle p_b,\varepsilon,A\rangle,\langle p_\varepsilon,\varepsilon\rangle\rangle,\\
&\langle\langle p_{ba},\varepsilon,A\rangle,\langle p_a,\varepsilon\rangle\rangle,\ \langle\langle p_{bba},\varepsilon,A\rangle,\langle p_{ba},\varepsilon\rangle\rangle\}.
\end{aligned}$$



Задачи

6.2.1. Пусть гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b\}^*$ задан соотношениями $h(a) = b$, $h(b) = \varepsilon$, $h(c) = a$. Рассмотрим язык L, порождаемый грамматикой $S \to a$, $S \to bS$, $S \to cSS$. Найти контекстно-свободную грамматику для языка $h(L)$.

6.2.2. Пусть гомоморфизм $h\colon \{a,b,c,d\}^* \to \{a,b\}^*$ задан соотношениями $h(a) = ba$, $h(b) = b$, $h(c) = a$, $h(d) = ab$. Рассмотрим язык L, порождаемый грамматикой $S \to SaSbcS$, $S \to d$. Является ли язык $h(L)$ праволинейным?

6.2.3. Рассмотрим гомоморфизм $h\colon \{a,b\}^* \to \{d,e,f\}^*$, заданный соотношениями $h(a) = fed$, $h(b) = ff$, и язык L, порождаемый грамматикой $R \to RaR$, $R \to TbT$, $T \to aT$, $T \to baR$, $T \to \varepsilon$. Найти контекстно-свободную грамматику, порождающую язык $h(L)$.

6.2.4. Рассмотрим гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b,c\}^*$, заданный соотношениями $h(a) = abab$, $h(b) = ababa$, $h(c) = ba$, и язык L, порождаемый грамматикой $R \to TaRac$, $R \to baT$, $T \to TabT$, $T \to ccR$, $T \to \varepsilon$. Найти контекстно-свободную грамматику, порождающую язык $h(L)$.

6.2.5. Пусть гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b\}^*$ задан соотношениями $h(a) = ab$, $h(b) = aaba$, $h(c) = b$. Рассмотрим язык L, порождаемый грамматикой $S \to aS$, $S \to aSbS$, $S \to bSaS$. Описать язык $h^{-1}(L)$.

6.2.6. Рассмотрим гомоморфизм $h\colon \{c,d,e\}^* \to \{a,b\}^*$, заданный соотношениями $h(c)=a$, $h(d)=ba$, $h(e)=bb$, и язык L, порождаемый грамматикой $B \to aBBa$, $B \to b$. Найти контекстно-свободную грамматику, порождающую язык $h^{-1}(L)$

6.2.7. Рассмотрим гомоморфизм $h\colon \{c,d,e\}^* \to \{a,b\}^*$, заданный соотношениями $h(c) = bba$, $h(d) = aa$, $h(e) = b$, и язык L, порождаемый грамматикой $S \to aSbS$, $S \to \varepsilon$. Найти контекстно-свободную грамматику, порождающую язык $h^{-1}(L)$

6.2.8. Рассмотрим гомоморфизм $h_1\colon\{a,b,c,d,e,f\}^* \to \{a,b,c\}^*$, заданный соотношениями $h_1(a) = ac$, $h_1(b) = bc$, $h_1(c) = aa$, $h_1(d) = ab$, $h_1(e) = ba$, $h_1(f) = bb$, гомоморфизм $h_2\colon \{a,b,c,d,e,f\}^* \to \{a,b\}^*$, заданный соотношениями $h_2(a) = a$, $h_2(b) = b$, $h_2(c) = ab$, $h_2(d) = aa$, $h_2(e) = bb$, $h_2(f) = ba$, и язык L, порождаемый грамматикой $S \to aSb$, $S \to ba$. Найти контекстно-свободную грамматику, порождающую язык $h_2(h_1^{-1}(L \cdot \{\varepsilon, c\}))$

6.2.9. Рассмотрим гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b\}^*$, заданный соотношениями $h(a) = aa$, $h(b) = ab$, $h(c) = ba$, и язык L, порождаемый грамматикой $T \to aTabT$, $T \to \varepsilon$. Найти контекстно-свободную грамматику, порождающую язык $h^{-1}(L)$.

6.2.10. Рассмотрим гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b,c\}^*$, заданный соотношениями $h(a) = abab$, $h(b) = ababa$, $h(c) = ba$, и язык L, порождаемый грамматикой $T \to TabTba$, $T \to aa$, $T \to \varepsilon$. Найти контекстно-свободную грамматику, порождающую язык $h^{-1}(L)$.

6.2.11. Рассмотрим гомоморфизм $h\colon \{a,b,c,d\}^* \to \{a,b\}^*$, заданный соотношениями $h(a) = aa$, $h(b) = ab$, $h(c) = ba$, $h(d) = bb$, и язык L, порождаемый грамматикой $S \to SabaSa$, $S \to abb$, $S \to \varepsilon$. Найти контекстно-свободную грамматику, порождающую язык $h^{-1}(L)$.

6.2.12. Пусть гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b,c\}^*$ задан соотношениями $h(a) = a$, $h(b) = cbc$, $h(c) = bc$. Рассмотрим язык $L = \{w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c\}$. Является ли язык $h^{-1}(L)$ контекстно-свободным?

6.2.13. Пусть гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b,c\}^*$ задан соотношениями $h(a) = a$, $h(b) = cac$, $h(c) = bc$. Рассмотрим язык $L = \{w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c\}$. Является ли язык

$h^{-1}(L)$ контекстно-свободным?

6.2.14. Пусть гомоморфизм $h\colon \{a,b,c\}^* \to \{a,b,c\}^*$ задан соотношениями $h(a) = a$, $h(b) = cac$, $h(c) = b$. Рассмотрим язык $L = \{w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c\}$. Является ли язык $h^{-1}(L)$ контекстно-свободным?

6.2.15. Существует ли такой контекстно-свободный язык L над алфавитом $\{a,b\}$, что язык $\{a^{|w|_b} \mid w \in L\}$ не является контекстно-свободным?

6.2.16. Существует ли такой контекстно-свободный язык L над алфавитом $\{a,b\}$, что язык $\{w \in \{a,b\}^* \mid a^{|w|_a} \in L\}$ не является контекстно-свободным?

6.2.17. Существует ли такой контекстно-свободный язык L над алфавитом $\{a,b\}$, что язык $\{w \in \{a,b\}^* \mid (\exists x \in L)\, |x| = |w|\}$ не является контекстно-свободным?

6.3. Представления контекстно-свободных языков посредством гомоморфизмов

[20, 11.3], [23, с. 103—109, 115], [15, с. 331—333], [7, 6.5], [11, 15.2]

Теорема 1. *Рассмотрим алфавит $\Sigma_0 = \{a_1, a_2, b_1, b_2, c\}$ и язык $L_0 \subseteq \Sigma_0^*$, порождаемый контекстно-свободной грамматикой G_0: $S \to Cb_1b_2b_1RC$, $C \to c$, $C \to cDC$, $D \to A$, $D \to AD$, $A \to a_1$, $A \to a_2$, $A \to b_1$, $A \to b_2$, $R \to \varepsilon$, $R \to Ta_1Tb_1$, $R \to Ta_2Tb_2$, $T \to R$, $T \to RCC$. Произвольный язык $L \subseteq \Sigma^*$ является контекстно-свободным тогда и только тогда, когда существует такой гомоморфизм $h\colon \Sigma^* \to \Sigma_0^*$, что $L = h^{-1}(L_0)$ или $L = h^{-1}(L_0 \cup \{\varepsilon\})$.*

Доказательство. Достаточность следует из теоремы 3 на с. 61. Приведём теперь идею доказательства необходимости (полное доказательство можно найти в [23, с. 103—109]).

Пусть дан произвольный контекстно-свободный язык L. Согласно теореме 5 на с. 38 язык $L - \{\varepsilon\}$ порождается некоторой контекстно-свободной грамматикой $\langle \{A_1, \ldots, A_n\}, \Sigma, P, A_1 \rangle$, в которой каждое правило имеет один из следующих трёх видов: $A_i \to a$, $A_i \to aA_j$, $A_i \to aA_jA_k$, где $a \in \Sigma$.

Определим вспомогательную функцию $\bar{h}$, ставящую в соответствие каждому символу из Σ конечный язык над алфавитом

Σ_0 следующим образом:

$$\begin{aligned}\bar{h}(a) = \{b_1b_2^ib_1 \mid (A_i \to a) \in P\} \cup \\ \cup \{b_1b_2^ib_1a_1a_2^ja_1 \mid (A_i \to aA_j) \in P\} \cup \\ \cup \{b_1b_2^ib_1a_1a_2^ka_1a_1a_2^ja_1 \mid (A_i \to aA_jA_k) \in P\}.\end{aligned}$$

Искомый гомоморфизм h определяется следующим образом: если

$$\bar{h}(a) = \{w_1, w_2, \dots, w_k\},$$

положим

$$h(a) = cw_1cw_2c \dots cw_kc.$$

□

Пример 2. Пусть $\Sigma = \{d, f, g\}$. Рассмотрим язык L, порождаемый грамматикой $A_1 \to f$, $A_1 \to dA_1A_2$, $A_2 \to fA_3$, $A_2 \to fA_2A_3$, $A_3 \to g$. Тогда $L = h^{-1}(L_0)$, где гомоморфизм h задан равенствами

$$\begin{aligned}h(d) &= cb_1b_2b_1a_1a_2a_2a_1a_1a_2a_1c, \\ h(f) &= cb_1b_2b_1cb_1b_2b_2b_1a_1a_2a_2a_2a_1cb_1b_2b_2b_1a_1a_2a_2a_2a_1a_1a_2a_2a_1c, \\ h(g) &= cb_1b_2b_2b_2b_1c.\end{aligned}$$

Рассмотрим, например, слово $dffg \in L$. Проверим, что слово $h(dffg)$ выводится в грамматике G_0 из теоремы 1. Очевидно, что $S \overset{*}{\Rightarrow} cb_1b_2b_1Rc$. С помощью последних пяти правил грамматики G_0 можно вывести, что

$$R \overset{*}{\Rightarrow} a_1a_2a_2a_1a_1a_2a_1CCb_1b_2b_1CCb_1b_2b_2b_1a_1a_2a_2a_2a_1CCb_1b_2b_2b_2b_1.$$

Осталось найти такие шесть выводимых из C слов $w_1, \dots, w_6$, что

$$\begin{aligned}&h(dffg) = \\ &= cb_1b_2b_1a_1a_2^2a_1a_1a_2a_1w_1w_2b_1b_2b_1w_3w_4b_1b_2^2b_1a_1a_2^3a_1w_5w_6b_1b_2^3b_1c.\end{aligned}$$

Подходят слова

$$w_1 = c,$$

$w_2 = c,$

$w_3 = cb_1b_2b_2b_1a_1a_2a_2a_2a_1cb_1b_2b_2b_1a_1a_2a_2a_2a_1a_1a_2a_2a_1c,$

$w_4 = cb_1b_2b_1c,$

$w_5 = cb_1b_2b_2b_1a_1a_2a_2a_2a_1a_1a_2a_2a_1c,$

$w_6 = c.$

Теорема 3 (Теорема Хомского—Шютценберже)**.** *Язык $L \subseteq \Sigma^*$ является контекстно-свободным тогда и только тогда, когда существуют такие натуральное число n, автоматный язык L_1 над алфавитом $\Sigma_n^{(\mathrm{D})} = \{a_1, b_1, a_2, b_2, \ldots, a_n, b_n\}$ и гомоморфизм $h\colon \big(\Sigma_n^{(\mathrm{D})}\big)^* \to \Sigma^*$, что $L = h(L_n^{(\mathrm{D})} \cap L_1)$, где $L_n^{(\mathrm{D})}$ — язык Дика над $2n$ буквами.*

Доказательство можно найти в [15, с. 331—333].

Задачи

6.3.1. Описать язык над алфавитом $\{a_1, a_2, b_1, b_2, c\}$, порождаемый грамматикой $C \to c$, $C \to cDC$, $D \to A$, $D \to AD$, $A \to a_1$, $A \to a_2$, $A \to b_1$, $A \to b_2$.

6.3.2. Описать язык над алфавитом $\{a_1, a_2, b_1, b_2, c\}$, порождаемый грамматикой $K \to CC$, $C \to c$, $C \to cDC$, $D \to A$, $D \to AD$, $A \to a_1$, $A \to a_2$, $A \to b_1$, $A \to b_2$.

6.3.3. Обозначим через L язык над алфавитом $\{d, f\}$, порождаемый грамматикой $J \to fJf$, $J \to dK$, $K \to dKd$, $K \to f$. Найти такой гомоморфизм $h\colon \{d, f\}^* \to \{a_1, a_2, b_1, b_2, c\}^*$, что $L = h^{-1}(L_0)$ (язык L_0 определён в теореме 1).

6.3.4. Обозначим через L язык над алфавитом $\{d_1, d_2, d_3, d_4\}$, порождаемый грамматикой $S \to d_1Sd_2Sd_3$, $S \to d_4$. Найти такой гомоморфизм $h\colon \{d_1, d_2, d_3, d_4\}^* \to \{a_1, a_2, b_1, b_2, c\}^*$, что $L = h^{-1}(L_0)$ (язык L_0 определён в теореме 1).

6.3.5. Пусть $\Sigma = \{d, f\}$. Обозначим через L язык, порождаемый грамматикой $J \to ffJd$, $J \to d$. Найти такой гомоморфизм $h\colon \{d, f\}^* \to \{a_1, a_2, b_1, b_2, c\}^*$, что $L = h^{-1}(L_0)$ (язык L_0 определён в теореме 1).

6.3.6. Рассмотрим язык L_1, порождаемый грамматикой $S \to aSbS$, $S \to \varepsilon$, и язык L_2, порождаемый грамматикой $S \to bbSa$, $S \to \varepsilon$. Найти такой гомоморфизм

$h\colon \{a, b\}^* \to \{a, b\}^*$, что

$$h(L_1 \cap \{a^m b^n \mid m \geqslant 0,\ n \geqslant 0\}) = L_2.$$

7. Детерминированные контекстно-свободные языки

7.1. Детерминированные автоматы с магазинной памятью

[20, 12.1], [2, 2.5.4], [27, 6.4], [9, с. 423], [24, 4.1], [25, с. 38], [22, с. 88—89], [19, 6.4]

Определение 1. Будем говорить, что два перехода МП-автомата $\langle\langle p_1, x_1, \beta_1\rangle, \langle q_1, \gamma_1\rangle\rangle$ и $\langle\langle p_2, x_2, \beta_2\rangle, \langle q_2, \gamma_2\rangle\rangle$ являются *совместными*, если

1) $p_1 = p_2$;

2) $x_1 \sqsubset x_2$ или $x_2 \sqsubset x_1$;

3) $\beta_1 \sqsubset \beta_2$ или $\beta_2 \sqsubset \beta_1$.

Определение 2. МП-автомат называется *детерминированным* (deterministic), если он имеет ровно одно начальное состояние и все переходы этого автомата попарно несовместны.

Пример 3. МП-автомат M из примера 5 на с. 58 не является детерминированным, так как переходы $\langle\langle 2, a, \varepsilon\rangle, \langle 1, D\rangle\rangle$ и $\langle\langle 2, \varepsilon, D\rangle, \langle 2, \varepsilon\rangle\rangle$ совместны.

Определение 4. Язык L над алфавитом Σ называется *детерминированным контекстно-свободным языком*, если существует детерминированный МП-автомат, распознающий язык $L \cdot \{\$\}$ над алфавитом $\Sigma \cup \{\$\}$, где **\$** — дополнительный символ, не принадлежащий множеству Σ. Символ **\$** называется *маркером конца строки.*

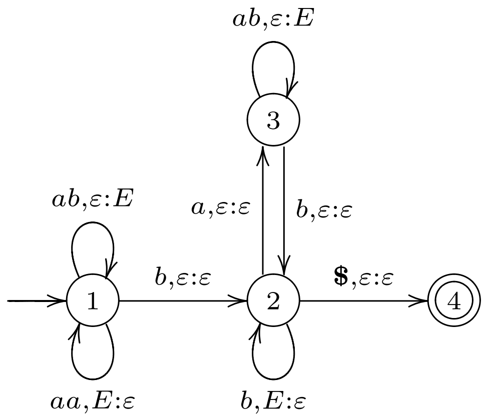
Пример 5. Рассмотрим алфавит $\Sigma = \{a, b\}$. Язык $L = \{a^m b^n \mid m = n \text{ или } n = 0\}$ — детерминированный контекстно-свободный язык, так как язык $L \cdot \{\$\}$ порождается детерминированным МП-автоматом (хотя язык L не порождается никаким детерминированным МП-автоматом).

Пример 6. Язык L, распознаваемый МП-автоматом M из примера 5 на с. 58, является детерминированным контекстно-свободным языком, так как язык $L \cdot \{\$\}$ порождается детерминированным МП-автоматом

$$M' = \langle \{1,2,3,4\}, \{a,b,\$\}, \{E\}, \Delta', \{1\}, \{4\} \rangle,$$

где

$$\begin{aligned}\Delta' = \{&\langle\langle 1, ab, \varepsilon\rangle, \langle 1, E\rangle\rangle,\ \langle\langle 1, aa, E\rangle, \langle 1, \varepsilon\rangle\rangle,\ \langle\langle 1, b, \varepsilon\rangle, \langle 2, \varepsilon\rangle\rangle,\\ &\langle\langle 2, b, E\rangle, \langle 2, \varepsilon\rangle\rangle,\ \langle\langle 2, a, \varepsilon\rangle, \langle 3, \varepsilon\rangle\rangle,\ \langle\langle 2, \$, \varepsilon\rangle, \langle 4, \varepsilon\rangle\rangle,\\ &\langle\langle 3, ab, \varepsilon\rangle, \langle 3, E\rangle\rangle,\ \langle\langle 3, b, \varepsilon\rangle, \langle 2, \varepsilon\rangle\rangle\}.\end{aligned}$$



Задачи

7.1.1. Является ли детерминированным контекстно-свободный язык $\{a^n b^{2m} a b^m a^n \mid m \geqslant 0,\ n \geqslant 0\}$?

7.1.2. Является ли детерминированным контекстно-свободный язык $\{a^k b^m c^n \mid k < \max(m,n)\}$?

7.1.3. Является ли детерминированным контекстно-свободный язык $\{w \in \{a,b\}^* \mid (|w|_a - |w|_b) \,\vdots\, 3\}$?

7.1.4. Является ли детерминированным контекстно-свободный язык $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$?

7.1.5. Является ли детерминированным контекстно-свободный язык $\{w \in \{a,b\}^* \mid |w|_a = |w|_b \geqslant 3\}$?

7.1.6. Является ли детерминированным контекстно-свободный язык $\{w \in \{a,b\}^* \mid |w|_a \geqslant |w|_b\}$?

7.1.7. Является ли детерминированным контекстно-свободный язык $\{w \in \{a,b\}^* \mid |w|_a \geqslant |w|_b$ и $|w|_a \geqslant 3\}$?

7.1.8. Является ли детерминированным контекстно-свободный язык $\{w \in \{a, b, c\}^* \mid |w|_a \geqslant |w|_b \text{ и } |w|_c \geqslant 3\}$?

7.1.9. Является ли детерминированным контекстно-свободный язык $\{w \in \{a, b, c\}^* \mid |w|_a \leqslant |w|_b \leqslant 3\}$?

7.1.10. Является ли детерминированным контекстно-свободный язык $\{w \in \{a, b, c\}^* \mid |w|_a \leqslant |w|_b \leqslant \min(3, |w|_c)\}$?

7.2. Свойства класса детерминированных контекстно-свободных языков

[20, 12.2], [2, 2.5.4, 2.6.4], [22, с. 90—93], [27, 6.4], [9, с. 424—425], [31, с. 160—161]

Теорема 1. *Каждый автоматный язык является детерминированным контекстно-свободным языком.*

Доказательство. Утверждение следует из теоремы 18 на с. 21. □

Теорема 2. *Пусть L — детерминированный контекстно-свободный язык. Тогда язык L не является существенно неоднозначным.*

Теорема 3. *Дополнение каждого детерминированного контекстно-свободного языка является детерминированным контекстно-свободным языком.*

Доказательство можно найти в [6, с. 110—116] или [2, с. 212—217].

Пример 4. Язык $L = \{a^k b^m c^n \mid k \neq m \text{ или } m \neq n\}$ над алфавитом $\{a, b, c\}$ не является детерминированным контекстно-свободным языком, так как его дополнение не является контекстно-свободным.

Задачи

7.2.1. Является ли детерминированным контекстно-свободный язык $\{a^k b^m c^n \mid k < \max(m, n)\}$?

7.2.2. Является ли детерминированным контекстно-свободный язык $\{v_1 v_2 v_2 v_3 \mid v_1 \in \{a, b\}^+,\ v_2 \in \{a, b\}^+,\ v_3 \in \{a, b\}^+\}$?

7.2.3. Является ли детерминированным контекстно-свободный язык $\{(abc)^k bc(bac)^m cb(acb)^n \mid k \neq m \vee m \neq n\}$?

7.2.4. Является ли детерминированным контекстно-свободный язык $\{(abc)^k bc(cba)^m cb(acb)^n \mid k \neq m \vee m \neq n\}$?

7.2.5. Является ли детерминированным контекстно-свободный язык $\{(abc)^k bc(cba)^m cb(acb)^n \mid k = m \vee m \neq n\}$?

7.2.6. Является ли детерминированным контекстно-свободный язык $\{(abc)^k bc(cba)^m cb(acb)^n \mid k = m \vee m = n\}$?

7.2.7. Является ли детерминированным контекстно-свободный язык, порождаемый грамматиклй $S \to ab$, $S \to SaaSbS$?

7.2.8. Является ли детерминированным контекстно-свободный язык, порождаемый грамматиклй $S \to aRb$, $R \to \varepsilon$, $R \to TT$, $T \to aSb$?

7.2.9. Является ли контекстно-свободным язык $\{a, b\}^* - L$, где язык L порождается грамматикой $S \to \varepsilon$, $S \to aSbS$?

Список литературы

[1] *Ахо А., Сети Р., Ульман Дж.* Компиляторы: принципы, технологии и инструменты. — М.: Вильямс, 2001. — 768 с.

[2] *Ахо А., Ульман Дж.* Теория синтаксического анализа, перевода и компиляции. Т. 1: Синтаксический анализ. — М.: Мир, 1978. — 612 с.

[3] *Ахо А., Хопкрофт Дж., Ульман Дж.* Построение и анализ вычислительных алгоритмов. — М.: Мир, 1979. — 536 с.

[4] *Братчиков И. Л.* Синтаксис языков программирования. — М.: Наука, 1975. — 232 с.

[5] *Волкова И. А., Руденко Т. В.* Формальные грамматики и языки. Элементы теории трансляции. 2-е изд., перераб. и доп. — М.: Диалог-МГУ, 1999. — 62 с.

[6] *Гинзбург С.* Математическая теория контекстно-свободных языков. — М.: Мир, 1970. — 326 с.

[7] *Гладкий А. В.* Формальные грамматики и языки. — М.: Наука, 1973. — 368 с.

[8] *Гладкий А. В., Мельчук И. А.* Элементы математической лингвистики. — М.: Наука, 1969. — 192 с.

[9] *Гордеев А. В., Молчанов А. Ю.* Системное программное обеспечение. — СПб.: Питер, 2001. — 736 с.

[10] *Грис Д.* Конструирование компиляторов для цифровых вычислительных машин. — М.: Мир, 1975. — 544 с.

[11] *Гросс М., Лантен А.* Теория формальных грамматик. — М.: Мир, 1971. — 294 с.

[12] *Карпов Ю. Г.* Теория автоматов. — СПб.: Питер, 2002. — 224 с.

[13] *Компаниец Р. И., Маньков Е. В., Филатов Н. Е.* Системное программирование. Основы построения трансляторов. — СПб.: КОРОНА принт, 2000. — 256 с.

[14] *Кук Д., Бейз Г.* Компьютерная математика. — М.: Наука, 1990. — 384 с.

[15] *Лаллеман Ж.* Полугруппы и комбинаторные приложения. — М.: Мир, 1985. — 440 с.

[16] Логический подход к искусственному интеллекту: От классической логики к логическому программированию / А. Тейз, П. Грибомон, Ж. Луи и др. — М.: Мир, 1990. — 432 с.

[17] *Льюис Ф., Розенкранц Д., Стирнз Р.* Теоретические основы проектирования компиляторов. — М.: Мир, 1979. — 656 с.

[18] *Мартыненко Б. К.* Языки и трансляции. — СПб.: Издательство С.-Петербургского университета, 2004. — 235 с.

[19] *Мозговой М. В.* Классика программирования: алгоритмы, языки, автоматы, компиляторы. Практический подход. — СПб.: Наука и Техника, 2006. — 320 с.

[20] *Пентус А. Е., Пентус М. Р.* Математическая теория формальных языков. — М.: Интернет-университет информационных технологий; БИНОМ. Лаборатория знаний, 2006. — 247 с. —(Серия «Основы информатики и математики»).

[21] *Пентус А. Е., Пентус М. Р.* Конечные автоматы и регулярные выражения. Сборник задач. — М.: Издательство попечительского совета механико-математического факультета МГУ, 2015. — 72 с.

[22] *Рейуорд-Смит В. Дж.* Теория формальных языков. Вводный курс. — М.: Радио и связь, 1988. — 128 с.

[23] *Саломаа А.* Жемчужины теории формальных языков. — М.: Мир, 1986. — 159 с.

[24] *Серебряков В. А., Галочкин М. П., Гончар Д. Р., Фуругян М. Г.* Теория и реализация языков программирования. — М.: МЗ-Пресс, 2003. — 345 с.

[25] *Соколов В. А., Кушниренко О. Б., Бадин Н. М.* Формальные языки и грамматики. Задачи и упражнения. — Ярославль: Ярославский государственный университет, 1993. — 55 с.

[26] *Трахтенброт Б. А., Барздинь Я. М.* Конечные автоматы (поведение и синтез). — М.: Наука, 1970. — 400 с.

[27] *Хопкрофт Дж. Э., Мотвани Р., Ульман Дж. Д.* Введение в теорию автоматов, языков и вычислений, 2-е изд. — М.: Вильямс, 2002. — 528 с.

[28] *Шень А.* Программирование: теоремы и задачи. — М.: МЦНМО, 1995. — 264 с.

[29] *Kinber E., Smith C.* Theory of Computing: a Gentle Introduction. — Prentice Hall, 2001. — 207 p.

[30] *Lewis H. R., Papadimitriou C. H.* Elements of the Theory of Computation. — Prentice Hall, 1981. — 466 p.

[31] *Lewis H. R., Papadimitriou C. H.* Elements of the Theory of Computation. 2nd ed. — Prentice Hall, 1998. — 361 p.

[32] *Sipser M.* Introduction to the Theory of Computation. — PWS Publishing company, 1997. — 396 p.