Регулярные выражения

Задача 1. Задайте регулярными выражениями следующие множества слов:
а) слова в алфавите $\{a, b\}$, такие что на третьем месте от начала слова стоит буква a, а на пятом месте от конца — буква b;
б) слова в алфавите $\{a, b\}$, в которых число букв a чётно;
в) слова в алфавите $\{a, b\}$, не содержащие подстроки ab;
г) слова в алфавите $\{a, b\}$, не содержащие подстроки abb;
д) слова в алфавите $\{a, b, c\}$, в которых нет двух соседних букв b;
е) слова в алфавите $\{a, b, c\}$, содержащие подслово вида bxa, где x — произвольная буква алфавита;
ё) слова в алфавите $\{a, b, c\}$, в которых за буквой a обязательно следует буква c;
ж) слова в алфавите $\{a, b, c\}$, в которых за буквой a не может идти буква b;
з) константы типа `double` языка Си (например, `1`, `9.86`, `.6`, `6.2e-23`).

Задача 2. Напишите регулярные выражения, задающие следующие языки над алфавитом $\{a, b\}$: **а)** $\{w \mid |w| \leqslant 3\}$; **б)** $\{w \mid |w| \geqslant 4\}$; **в)** $\{w \mid |w|_b \geqslant 2\}$; **г)** $\{w \mid |w| = 2$ или $|w|_a = 3\}$; **д)** $\{w \mid |w| = 3$ или $|w|_a \geqslant 5\}$.

Задача 3. Опишите множества слов, задаваемые следующими регулярными выражениями: **а)** $(a + c)^*$; **б)** $(a + b)^* a$; **в)** $b(b + c)^*$; **г)** $a(a + b)^* a$; **д)** $a^* b a^* b a^* b a^*$; **е)** $(a + b)^* b(a + b)(a + b)$; **ё)** $b^*(a + (ab^+)^*)$.

Задача 4. Является ли язык $\{a^n \mid$ существует такое $p \geqslant n$, что p простое и $p + 2$ простое$\}$ регулярным?

Задача 5. Докажите следующие равенства: **а)** $(1 + e + ee + \ldots + e^{n-1})(e^n)^* = e^*$ для любого $n \geqslant 1$; **б)** $(e^* f)^* e^* = (e + f)^*$; **в)** $1 + e(fe)^* f = (ef)^*$.

Задача 6. Упростите регулярные выражения: **а)** $(a + b + ab)^*$; **б)** $(a^* b)^* + (b^* a)^*$; **в)** $(abbaab + abbaaba)^*$; **г)** $(a + b)^* ab(a + b)^* + (a + b)^* a + b^*$; **д)** $1 + aa^* + bb^*$; **е)** $(a + b)^*(ab + ba)(a + b)^* + a^* + b^*$.

Задача 7. Докажите, что если регулярные выражения e, f и g таковы, что $e = ef + g$ и $\varepsilon \notin \mathcal{L}(f)$, то $e = gf^*$.

Задача 8. Существует ли такой язык L, что $L^* \neq \{x^n \mid x \in L, n \geqslant 0\}$?

Задача 9. **а)** Докажите, что для любых двух языков L_1 и L_2 выполнено равенство $(L_1 \cdot L_2)^R = L_2^R \cdot L_1^R$ (язык L^R состоит из слов языка L, написанных задом наперёд). **б)** Существует ли такой язык L, что $(L^R)^* \neq (L^*)^R$?

Задача 10. Докажите, что если язык L задается регулярным выражением, то и язык L^R тоже задается регулярным выражением.

Недетерминированные конечные автоматы

Задача 1. Найдите НКА, распознающие следующие множества слов:
а) слова в алфавите $\{a,b\}$, такие что на третьем месте от начала слова стоит буква a, а на пятом месте от конца — буква b;
б) слова в алфавите $\{a,b\}$, в которых число букв a чётно;
в) слова в алфавите $\{a,b\}$, не содержащие подстроки ab;
г) слова в алфавите $\{a,b\}$, не содержащие подстроки abb;
д) слова в алфавите $\{a,b\}$, в которых число букв a чётно, а число букв b нечётно;
е) слова в алфавите $\{a,b,c\}$, в которых нет двух соседних букв b;
ё) слова в алфавите $\{a,b,c\}$, содержащие подслово вида bxa, где x — произвольная буква алфавита;
ж) слова в алфавите $\{a,b,c\}$, в которых за буквой a обязательно следует буква c;
з) слова в алфавите $\{a,b,c\}$, в которых за буквой a не может идти буква b;
и) слова из букв a и b, такие что разность числа букв a и числа букв b делится на 3;
й) константы типа `double` языка Си (например, `1`, `9.86`, `.6`, `6.2e-23`). **к)** числа в римской системе счисления.

Задача 2. Найдите НКА, распознающие следующие языки над алфавитом $\{a,b\}$: **а)** $\{w \mid |w| \leqslant 3\}$; **б)** $\{w \mid |w| \geqslant 4\}$; **в)** $\{w \mid |w|_b \geqslant 2\}$; **г)** $\{w \mid |w| = 2$ или $|w|_a = 3\}$; **д)** $\{w \mid |w| = 3$ или $|w|_a \geqslant 5\}$. **е)** $\{w \mid |w|_a \geqslant 2$ или $|w|_b \geqslant 4\}$. **ё)** $\{w \mid |w|_a \leqslant 2$ и $|w| \leqslant 5\}$. **ж)** $\{w \mid |w|_a \leqslant 4$ и $|w|_b \geqslant 2\}$.

Задача 3. Найдите НКА, распознающие следующие множества слов:
а) $\{a, bb\}^* \quad (\Sigma = \{a,b\})$;
б) $\{a,b\}^+ - a^* - b^* \quad (\Sigma = \{a,b\})$;
в) $\{bc^{2n}ac^{2m} \mid m,n \geqslant 0\} \cup \{bc^{2n+2} \mid n \geqslant 0\} \quad (\Sigma = \{a,b,c\})$;
г) $\{(ab)^{2n+2}(aba)^{3m+1} \mid m,n \geqslant 0\} \quad (\Sigma = \{a,b\})$;
д) $\{(ab)^k(aab)^l \mid k,l > 0\} \quad (\Sigma = \{a,b\})$.

Задача 4. Найдите ДКА для языков, указанных в задачах 1-3.

Задача 5. Найдите ДКА, распознающие множества слов, задаваемые следующими регулярными выражениями:
а) $(a+b)^*b(a+1)b(a+b)^*$;
б) $a^*ba^+ba^*(ba^*+1)$;
в) $a((ab+ba)^*b)^*$;
г) $a(a(ab)^*b)^*+ba$.

Задача 6. Докажите, что если L — автоматный язык, то
а) Язык $Pref(L) = \{u \mid \exists w \in L(u -$ префикс $w\}$ автоматен;
б) Язык $Suff(L) = \{u \mid \exists w \in L(u -$ суффикс $w\}$ автомате;
в) Язык $Subw(L) = \{u \mid \exists w \in L(u -$ подслово $w\}$ автоматен;
г) Язык $Subseq(L) = \{u \mid \exists w \in L(u -$ подпоследовательность $w\}$ автоматен.

Задача 7. Докажите, что класс языков, распознаваемых НКА с одним начальным и одним завершающим состоянием, не совпадает с классом всех автоматных языков.

Задача 8.
а) Верно ли утверждение предыдущей задачи для класса языков без пустого слова?
б) Можно ли в определении НКА обойтись двумя завершающими состояниями?

Контекстно-свободные грамматики

Задача 1. Опишите языки, порождаемые следующими грамматиками:
а) $S \to FF$, $F \to FF$, $F \to ab$ **б)** $S \to FS$, $S \to FF$, $F \to aFb$, $F \to \varepsilon$
в) $F \to a$, $F \to bF$, $F \to cFF$ **г)** $F \to ab$, $F \to aFb$, $F \to FF$
д) $F \to a$, $F \to bF$, $F \to cFF$, $F \to dFFF$ **е)** $S \to \varepsilon$, $S \to aSSa$, $S \to bSSb$
ё) $S \to \varepsilon$, $S \to aSaSbS$

Задача 2. Постройте контекстно-свободные грамматики, задающие следующие языки:
а) $\{w \in \{a,b\}^* \mid w = w^R\}$ **б)** $\{a^n b^m \mid n > m\}$ **в)** $\{a^k b^{k+2} b^2 \mid k \in \mathbb{N}\}$ **г)** $\{a^n b^m \mid n - m \vdots 3\}$
д) $\{a^n b^{2n+1} \mid n \in \mathbb{N}\}$ **е)** $\{a^k b^{k+m} a^m \mid k, m \in \mathbb{N}\}$

Задача 3. Постройте грамматики в нормальной форме Хомского, эквивалентные следующим грамматикам:
а) $S \to aFbF, F \to aFb, F \to \varepsilon, F \to Ga$ **б)** $F \to a, F \to bF, F \to cFF$
в) $S \to TaT, T \to UV, U \to XY, X \to Xa, X \to \varepsilon, T \to XbX, X \to TXT$.

Задача 4. Перечислите все деревья вывода в грамматике $S \to a$, $S \to bR$, $R \to cTT$, $T \to a$, $T \to ba$.

Задача 5. Дана контекстно-свободная грамматика $S \to S \cdot F$, $S \to S + F$, $F \to (S)$, $S \to F$, $F \to 2$ в алфавите $\Sigma = \{2, +, \cdot,), (\}$. Применив алгоритм CYK, проверьте, принадлежат ли следующие слова языку, порождаемому этой грамматикой: **а)** $2\,2 + 2 \cdot 2$; **б)** $(2+2) \cdot 2$; **в)** $2 + 2 + 2$; **г)** $2 + 2 \cdot 2$. Для тех слов, которые принадлежат языку, постройте соответствующие деревья вывода. Для каких из них получится ровно одно дерево?

Задача 6. Постройте однозначную контекстно-свободную грамматику, порождающую тот же язык, что и грамматика **а)** $S \to SaS$, $S \to SbS$, $S \to c$, $S \to eSf$; **б)** $S \to aSaaaa$, $S \to aSaa$, $S \to aaSa$, $S \to b$.

Задача 7. Постройте однозначную контекстно-свободную грамматику для языка **а)** $\{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$; **б)** $\{w \in \{a,b\}^* \mid |w|_a \geqslant |w|_b\}$.

Задача 8. Постройте контекстно-свободную грамматику с двумя правилами, эквивалентную грамматике $S \to \varepsilon$, $S \to F$, $F \to ab$, $F \to aFb$, $F \to FF$.

Контекстно-свободные языки. Часть II

Задача 1. Постройте контекстно-свободную грамматику для языка $\{w \in \{a,b\}^* \mid |w|_a > |w|_b + 7\}$.

Задача 2. Является ли контекстно-свободным язык:

а) $\{www \mid w \in \{a,b\}^*\}$;
б) $\{uvvw \mid u,v,w \in \{a,b\}^+\}$;
в) $\{a^i b^j c^k \mid 0 \leqslant i \leqslant j \leqslant k\}$;
г) $\{a^k b^l a^m b^n \mid k + l = m + n\}$;
д) $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c\}$;
е) $\{u \in \{a,b\}^+ \mid u = u^R\}$;
ё) $\{uv \mid u,v \in \{a,b\}^+, u = u^R, v = v^R\}$;
ж) $\{uu \mid u \in \{a,b\}^*\}$;
з) $\{a,b\}^* - \{uu \mid u \in \{a,b\}^*\}$;
и) $\{w \in \{a,b,c\}^* \mid |w|_a \neq |w|_b$ или $|w|_b \neq |w|_c\}$?

Задача 3. Существуют ли такие контекстно-свободные языки $L_1 \subseteq \{a,b\}^*$ и $L_2 \subseteq \{b,c\}^*$, что **а)** язык $L_1 \cap L_2$ не является автоматным; **б)** язык $L_1 - L_2$ не является контекстно-свободным; **в)** язык $L_1 - L_2$ не является автоматным?

Задача 4. Однозначна ли контекстно-свободная грамматика **а)** $E \to EcE$, $E \to EdE$, $E \to aEb$, $E \to f$; **б)** $K \to iUtKeKf$, $K \to iUtKf$, $K \to p$, $U \to a$, $U \to b$; **в)** $S \to TV$, $S \to VT$, $T \to aVa$, $T \to bVb$, $V \to aTa$, $V \to bTb$, $T \to \varepsilon$; **г)** $S \to a$, $S \to b$, $S \to cSS$, $S \to dSS$?

Задача 5. Существует ли однозначная контекстно-свободная грамматика для языка $\{a^k b^m \mid 1 \leqslant k \leqslant 2m\}$?

Задача 6. Существует ли над алфавитом $\{a,b\}$ такой язык контекстно-свободный язык L, что язык $\{ubv \mid u \in L$ и $uv \in L\}$ не является контекстно-свободным?

МП-автоматы

Задача 1. Постройте МП-автоматы (где возможно — детермированные МП-автоматы) для следующих языков: **а)** $\{a^n b^m \mid n \neq m\}$; **б)** $\{w \in \{a,b\}^* \mid |w|_a = 2|w|_b\}$; **в)** $\{w \in \{a,b,c\}^* \mid |w|_a = |w|_b$ или $|w|_b = |w|_c\}$; **г)** множество арифметических выражений над переменными x и y; **д)** множество регулярных выражений над алфавитом $\{a,b,c\}$.

Задача 2. Докажите, что пересечение контекстно-свободного языка и регулярного языка является контекстно-свободным языком.

Задача 3. Относительно каких операций замкнут класс детерминированных контекстно-свободных языков (т. е. языков, задаваемых детерминированными МП-автоматами): **а)** пересечение; **б)** объединение; **в)** дополнение; **г)** звёздочка Клини?

Задача 4. Является ли детерминированным контекстно-свободным языком язык $\{a^k b^m c^n \mid k < \max\{m,n\}\}$?

Задача 5. Существует ли язык, для которого есть однозначная контекстно-свободная грамматика, но нет детерминированного МП-автомата?

Конечные преобразователи и задаваемые ими преобразования

Определение. *Конечным преобразователем* называется кортеж $M = \langle Q, \Sigma, \Sigma', \Delta, q_0, F\rangle$, где Q — конечное множество состояний, Σ, Σ' — входной и выходной алфавиты соответственно, а Δ задаёт на множестве Q структуру ориентированного графа, дуги которого помечены парами $\langle x, \beta\rangle$, где $x \in \Sigma^*$, $\beta \in \Sigma'^*$. $q_0 \in Q$ — стартовое состояние, $F \subseteq Q$ — множество завершающих состояний. Слова $u \in \Sigma^*$ и $\alpha \in \Sigma'^*$ находятся в отношении R_M, если существует успешный (начинающийся в q_0 и заканчивающийся в одном из состояний из F) путь, на котором из первых компонент меток складывается слово u, из вторых — α. Для языка $L \subseteq \Sigma^*$ положим $R_M(L) = \{\alpha \in \Sigma'^* \mid (\exists u \in L)\, uR_M\alpha\}$. Преобразования языков вида $L \mapsto R_M(L)$ называются *конечными преобразованиями.*

Предложение.

1. Композиция конечных преобразований есть конечное преобразование.
2. Конечное преобразование, применённое к регулярному языку, даёт регулярный язык.
3. Конечное преобразование, применённое к контекстно-свободному языку, даёт контекстно-свободный язык.

Задача 1. Докажите, что следующие отображения являются конечными преобразованиями:

1. $L \mapsto L \cap L_0$, где L_0 — регулярный язык;
2. $L \mapsto L \cup L_0$, где L_0 — регулярный язык;
3. $L \mapsto h^{-1}(L)$, где h — гомоморфизм.

Задача 2. Докажите, что следующие отношения задаются конечными преобразованиями (здесь x и y — натуральные числа, записанные в двоичной записи начиная *с младшего разряда*):

1. $x \to x + 1$;
2. $x \to x + k$ (k — фиксированное натуральное число);
3. $x \to \max(x - 1, 0)$;
4. $x \to \max(x - k, 0)$ (k — фиксированное натуральное число);
5. $x \to 2x + 1$;
6. $x \to kx$ (k — фиксированное натуральное число);
7. $(x, y) \to x+y$ (пара (x, y) записывается «по очереди»: младший разряд x, младший разряд y, следующий разряд x, следующий разряд y и т. д.).

Задача 3. а) Докажите, что если φ_1, φ_2 — конечные преобразования, то $L \mapsto \varphi_1(L) \cup \varphi_2(L)$ — тоже конечное преобразование.
б) Верно ли, что если φ_1, φ_2 — конечные преобразования, то $L \mapsto \varphi_1(L) \cap \varphi_2(L)$ — тоже конечное преобразование?

Задача 4. Задаются ли конечными преобразованиями следующие отображения?

1. $x \to x^2$, $x \in \mathbb{N}$.
2. умножение натуральных чисел.
3. Перекодировка из двоичной в унарную запись.
4. Перекодировка из двоичной в троичную запись.
5. $w \to w^R, w \in \Sigma^*$.