Декабрьская контрольная по ТРЯП задачи, решения и критерии

ФУПМ 2017

Разбалловка и общие положения

неуд	удовл	хорошо	отлично
$0 \leqslant \Sigma \leqslant 12$	$13 \leqslant \Sigma \leqslant 17$	$18 \leqslant \Sigma \leqslant 25$	$26 \leqslant \Sigma \leqslant 36$
1: 0-6, **2:** 7-12	**3:** 13-15, **4:** 16-17	**5:** 18-20, **6:** 21-23, **7:** 24-25	**8:** 26-30, **9:** 31-33, **10:** 34-36

В случае дробной суммы баллов, перед выставлением оценки происходит арифметическое округление.

Приведённые ниже критерии оценивания выработанны с учётом типовых ошибок и определяют общую политику проверки, однако заведомо не могут покрыть всевозможные случаи. При некритериальном случае, проверяющий оценивает решение исходя из здравого смысла и духа критериев. В случае несогласия с оценкой за работу, студент имеет право подать апелляцию. Апелляцию нужно подать в письменном виде во время показа работ (или в день показа до 23:59 только в случае невозможности это сделать во время показа). Жалобы и замечания, возникшие после показа работ можно направить письмом всем преподавателям курса (адреса сообщаются при желании подать таковые).

Внимание! подача апелляции приведёт к полному пересмотру работы апелляционной комиссией, в результате чего оценка может как повыситься, так и понизиться.

Напоминаем положения, указанные в преамбуле к контрольной.

1. Ответы, включая правильные, при отсутствии решений оцениваются в 0 (ноль) баллов.

2. Объекты, полученные «методом внимательного вглядывания», без доказательства корректности построения оцениваются в 0 (ноль) баллов.

3. При формулировке вопроса «верно ли, что», в случае положительного ответа приведите доказательство, а в случае отрицательного – контрпример. Верное рассуждение без контрпримера оценивается в половину задачи.

Задачи

Задача 1 (5).

1 $\langle$i$\rangle$. Заданы языки $L_1 = \{a^n b^{2n} | n \geqslant 0\}$ и $L_2 = \{a^n b^k | n \geqslant k \geqslant 1\}$ Постройте КС-грамматику и МП-автомат для языка $L_1(L_1 \cup L_2)^* L_2$.

1 $\langle$ii$\rangle$. Заданы языки $L_1 = \{a^n b^k | k \geqslant n \geqslant 0\}$ и $L_2 = \{a^{2n} b^n | n \geqslant 1\}$ Постройте КС-грамматику и МП-автомат для языка $((L_1 \cup L_2)L_1)^* L_2$.

1 $\langle$iii$\rangle$. Заданы языки $L_1 = \{b^n a^{2n} | n \geqslant 0\}$ и $L_2 = \{b^n a^k | n \geqslant k \geqslant 1\}$ Постройте КС-грамматику и МП-автомат для языка $(L_1(L_1 \cup L_2))^* L_2$.

1 $\langle$iv$\rangle$. Заданы языки $L_1 = \{b^n a^k | k \geqslant n \geqslant 0\}$ и $L_2 = \{b^{2n} a^n | n \geqslant 1\}$ Постройте КС-грамматику и МП-автомат для языка $L_2((L_1 \cup L_2)L_1)^*$.

Указание. Строим грамматики для L_1 и L_2, а дальше строим над ними грамматику для языка $L_1(L_1 \cup L_2)^* L_2$, например, используя подстановку в грамматику для $a(a|b)^*b$.

Решение. 1 $\langle$i$\rangle$. Язык L_1 порождается грамматикой G_1, заданной правилами $S \to aSbb \mid \varepsilon$, а L_2 – грамматикой G_2: $S \to ASb \mid Ab$; $A \to aA \mid a$. Докажем корректность построения L_2, корректность построения L_1 доказывается аналогично и проще. Заметим, что нетерминал A порождает язык a^+ (формально это следует из решения уравнения $A = aA + a$). Докажем, что $L(G_2) \subseteq L_2$. Заметим, что для каждого слова, выводимого из G_2 существует вывод, в котором сначала применяются правила для нетерминала S, а затем для нетерминала A. Пусть при выводе слова w в G_2 первое правило было применено m раз. Тогда вывод w устроен так: $S \overset{*}{\Rightarrow} A^m S b^m \Rightarrow A^{m+1} b^{m+1} \overset{*}{\Rightarrow} w$. Значит w имеет вид $a^k b^n$, где

$n = m+1$, а $k \geqslant n$ – последнее справедливо, поскольку из каждого нетерминала A выводятся только слова вида $a^t, t \geqslant 1$. Значит, мы показали, что $w \in L_2$. Докажем, что $L_2 \subseteq L(G_2)$. Достаточно для произвольного слова $a^n b^k, n \geqslant k \geqslant 1$ привести его вывод в грамматике G_2. Для этого применим $k-1$ раз первое правило, затем второе, после чего применим для некоторого A третье правило $n-k$ раз, а дальше применим для каждого A четвёртое правило.

Как хорошо известно из курса, КС-языки замкнуты относительно регулярных операций. Так, если G_A и G_B – грамматики с аксиомами S_A и S_B и множества нетерминалов грамматик не пересекаются, то для того, чтобы построить конкатенацию $L(G_A)L(G_B)$ достаточно объединить множества нетерминалов и правил двух грамматик, добавить новую аксиому S и добавить правило $S \to S_A S_B$. Для объединения нужно добавить правила $S \to S_A \mid S_B$, а для итерации $(L(G_A))^*$ – $S \to S_A S \mid \varepsilon$.

Занумеруем нетерминалы грамматики G_1 и G_2 и применив указанные преобразования получим искомую грамматику, заданную правилами:

$$S_A \to S_1 S_B \quad S_B \to S_C S_2 \quad S_C \to S_D S_C \mid \varepsilon \quad S_D \to S_1 \mid S_2$$
$$S_1 \to aS_1bb \mid \varepsilon \qquad S_2 \to A_2 S_2 b \mid A_2 b;\ A_2 \to aA_2 \mid a.$$

МП-автомат строим по алгоритму КС$\to$МП из курса.

Критерии.

- Построены грамматики $+\mathbf{1}$;
- Доказана корректность — $+\mathbf{1}$ за каждую сторону для трудной, $+\mathbf{0.5}$ за каждую сторону для простой. Если докзательство записано только для трудного случая и в нём нет помарок, то это даёт автоматом $+\mathbf{1}$ балл и для простого случая.
- Построена грамматика для требуемого языка $+\mathbf{1}$.
- Построен МП-автомат по соответствующей КС-грамматике $+\mathbf{1}$.

Задача 2 (б). **Задача 2** ⟨i⟩. Дана грамматика

$$G = \{\{A, S, B\}, \{a, b, c\}, \{S \to Aa; A \to a \mid Bc, B \to Sb\}, B\}.$$

1. Является ли грамматика G LR(k)-грамматикой? При положительном ответе на вопрос найти минимальное k и построить соответствующий анализатор.

2. Воспользовавшись **нижеследующей** управляющей таблицей LR(k)-анализатора, построить дерево разбора цепочки $ccabac\$$.

	Action				Goto								
I	c	$	a	b	S'	E	B	S	c	a	e	b	$
I_0	S	R (B→)				I_1	I_2		I_3				
I_1		A(S'→E$)											
I_2		R (E→B)											
I_3	S		R (S→)	R (B→)			I_5	I_4	I_6				
I_4			S							I_7			
I_5				S								I_8	
I_6	S		R (S→)	R (B→)			I_5	I_9	I_6				
I_7	S	R (B→cSa)							I_{10}				
I_8			R (S→Bb)										
I_9			S							I_{11}			
I_{10}		R (B→cSac)											
I_{11}	S			R (B→cSa)					I_{12}				
I_{12}				R (B→cSac)									

Задача 2 $\langle$ii$\rangle$**.** Дана грамматика

$$G = \{\{A, S, B\}, \{a, b, c\}, \{S \to Ac; A \to Ba \mid \varepsilon, B \to bS\}, B\}.$$

1. Является ли грамматика G LR(k)-грамматикой? При положительном ответе на вопрос найти минимальное k и построить соответствующий анализатор.

2. Воспользовавшись **нижеследующей** управляющей таблицей LR(k)-анализатора, построить дерево разбора цепочки $bcbaab\$$.

	Action				Goto								
I	$	b	c	a	S'	E	B	S	e	b	a	c	$
I_0	R (B→)	S				I_1	I_2			I_3			
I_1	A(S'→E$)												
I_2	R (E→B)												
I_3			S	R (S→)				I_4				I_5	
I_4				S							I_6		
I_5		S		R (B→)			I_7			I_8			
I_6	R (B→bSa)	S								I_9			
I_7				R (S→cB)									
I_8			S	R (S→)				I_{10}				I_5	
I_9	R (B→bSab)												
I_{10}				S							I_{11}		
I_{11}		S		R (B→bSa)						I_{12}			
I_{12}				R (B→bSab)									

Задача 2 $\langle$iii$\rangle$**.** Дана грамматика

$$G = \{\{A, S, B\}, \{a, b, c\}, \{S \to Ac; A \to Aa \mid B \mid \varepsilon, B \to bS\}, B\}.$$

1. Является ли грамматика G LR(k)-грамматикой? При положительном ответе на вопрос найти минимальное k и построить соответствующий анализатор.

2. Воспользовавшись **нижеследующей** управляющей таблицей LR(k)-анализатора, построить дерево разбора цепочки $ccacba\$$.

Action					Goto								
I	c	$	a	b	S'	E	B	S	c	a	e	b	$
I_0	S	R (B→)				I_1	I_2		I_3				
I_1		A(S'→E$)											
I_2		R (E→B)											
I_3	S		R (S→)	R (B→)			I_5	I_4	I_6				
I_4			S							I_7			
I_5				S								I_8	
I_6	S		R (S→)	R (B→)			I_5	I_9	I_6				
I_7	S	R (B→cSa)							I_{10}				
I_8			R (S→Bb)										
I_9			S							I_{11}			
I_{10}		R (B→cSac)											
I_{11}	S			R (B→cSa)					I_{12}				
I_{12}				R (B→cSac)									

Задача 2 ⟨iv⟩. Дана грамматика

$$G = \{\{A, S, B\}, \{a, b, c\}, \{S \to Ac; A \to Aa \mid a \mid B, B \to Sb\}, B\}.$$

1. Является ли грамматика G LR(k)-грамматикой? При положительном ответе на вопрос найти минимальное k и построить соответствующий анализатор.

2. Воспользовавшись **нижеследующей** управляющей таблицей LR(k)-анализатора, построить дерево разбора цепочки $bcbaba\$$.

Action					Goto								
I	$	b	c	a	S'	E	B	S	e	b	a	c	$
I_0	R (B→)	S				I_1	I_2			I_3			
I_1	A(S'→E$)												
I_2	R (E→B)												
I_3			S	R (S→)				I_4				I_5	
I_4				S							I_6		
I_5		S		R (B→)			I_7			I_8			
I_6	R (B→bSa)	S								I_9			
I_7				R (S→cB)									
I_8			S	R (S→)				I_{10}				I_5	
I_9	R (B→bSab)												
I_{10}				S							I_{11}		
I_{11}		S		R (B→bSa)						I_{12}			
I_{12}				R (B→bSab)									

Решение. Таблицы для LR-анализаторов можно получить введя на сайте http://lrk.umeta.ru правила грамматики в формате

S -> Aa
A -> a | Bc
B -> Sb

Критерии.

- Построена каноническая система множеств LR(1)-ситуаций — **2,5 балла**. (При неправильной аксиоме — **1,5 балла**). (При мелких описках — снижать **0,5 балла**).

- По канонической системе множеств правильно построена управляющая таблица LR(1) анализатора — **1,5 балла**. (Если по неправильному множеству строится правильная таблица — **0 баллов**)

- Доказательство не LR(0) — **1 балл**.

- Правлильная работа анализатора — **0,5 балла**. (При попытке использовать только что построенную таблицу – **0 баллов**).

- Правильное дерево — **0,5 балла**.

Задача 3 (2).

3 ⟨i⟩. Постройте LZW-автомат и SLG G_w для слова $w = abaababaabaab$.

Решение.

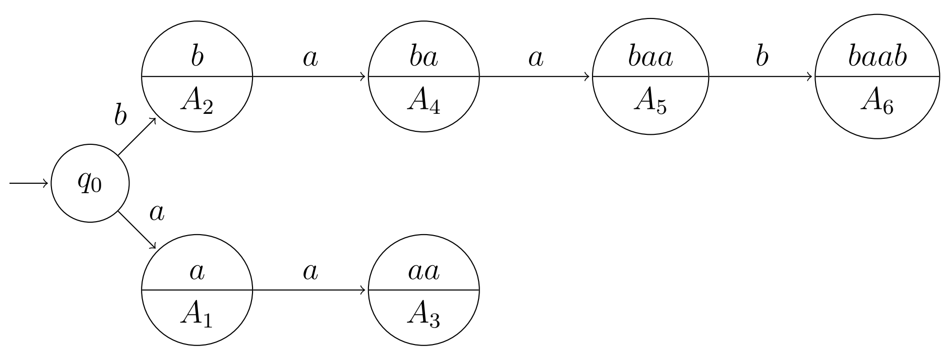


Рис. 1: LZW-автомат

По LZW-автомату построим грамматику по алгоритму:

$$S \to A_1A_2 \dots A_6, \quad A_1 \to a, \quad A_2 \to b, \quad A_3 \to A_1a,$$

$$A_4 \to A_2a, \quad A_5 \to A_4a, \quad A_6 \to A_5b.$$

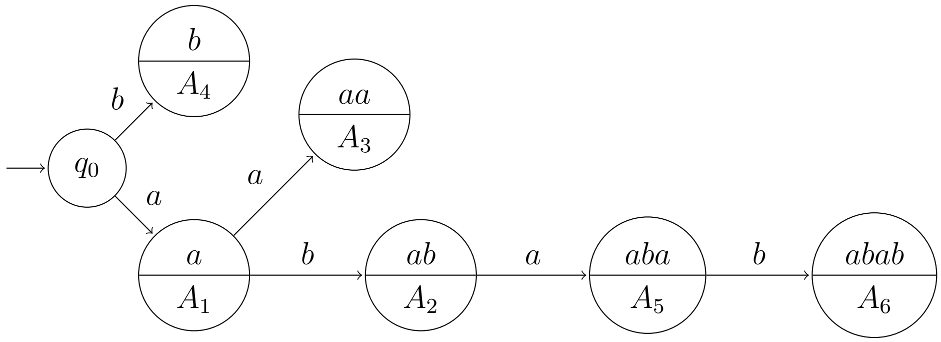


Рис. 2: LZW-автомат

3 $\langle$ii$\rangle$. Постройте LZW-автомат и SLG G_w для слова $w = aabaababaabab$.

Решение.

По LZW-автомату построим грамматику по алгоритму:

$$S \to A_1A_2 \dots A_6, \quad A_1 \to a, \quad A_2 \to A_1b, \quad A_3 \to A_1a,$$
$$A_4 \to b, \quad A_5 \to A_4a, \quad A_6 \to A_5b.$$

3 $\langle$iii$\rangle$. Постройте LZW-автомат и SLG G_w для слова $w = babbababbabba$.

Решение.

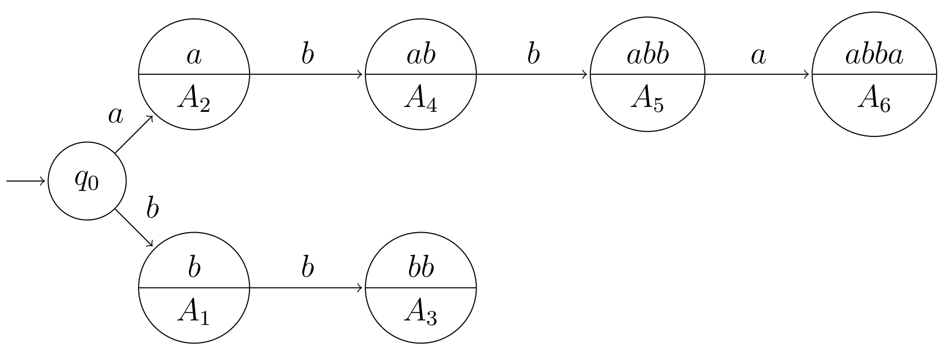


Рис. 3: LZW-автомат

По LZW-автомату построим грамматику по алгоритму:

$$S \to A_1A_2 \dots A_6, \quad A_1 \to b, \quad A_2 \to a, \quad A_3 \to A_1b,$$

$$A_4 \to A_2 b, \quad A_5 \to A_4 b, \quad A_6 \to A_5 a.$$

3 $\langle$iv$\rangle$. Постройте LZW-автомат и SLG G_w для слова $w = bbabbababbaba$.

Решение.

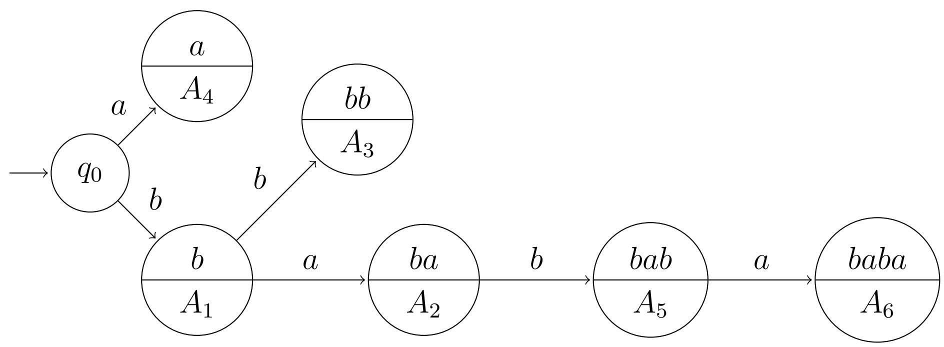


Рис. 4: LZW-автомат

По LZW-автомату построим грамматику по алгоритму:

$$S \to A_1 A_2 \dots A_6, \quad A_1 \to b, \quad A_2 \to A_1 a, \quad A_3 \to A_1 b,$$
$$A_4 \to a, \quad A_5 \to A_4 b, \quad A_6 \to A_5 a.$$

Критерии.

1.

- Разобраны только простые варианты разбиений (без пересечений): **1 балл**
- Полное решение, рассмотрены все разбиения, но есть ошибки: **2 балла**
- Полное решение, рассмотрены все разбиения: **4 балла**

2. Честно доказано, что пересечение L и регулярного языка R — это R и сказано, что регулярный язык является КС — **2 балла**. Без честного доказательства — **1 балл**.

Задача 4 (4+2).

Задача 4 $\langle$i$\rangle$**.** Является ли язык $L = \{a^k b^m c^n | k < \min(m, n)\} \subseteq \{a, b, c\}^*$ КС-языком? Является ли язык $(\Sigma^* \setminus L) \cap c^+(a \mid b)^*$ КС-языком?

Задача 4 $\langle$ii$\rangle$**.** Является ли язык $L = \{a^k b^m c^n | m \leqslant k, k < n\} \subseteq \{a, b, c\}^*$ КС-языком? Является ли язык $(\Sigma^* \setminus L) \cap c^* b^+ a^*$ КС-языком?

Задача 4 $\langle$iii$\rangle$**.** Является ли язык $L = \{a^k b^m c^n | k > \max(m, n)\} \subseteq \{a, b, c\}^*$ КС-языком? Является ли язык $(\Sigma^* \setminus L) \cap b^+(a \mid c)^*$ КС-языком?

Задача 4 $\langle$iv$\rangle$**.** Является ли язык $L = \{a^k b^m c^n | m \geqslant k, k < n\} \subseteq \{a, b, c\}^*$ КС-языком? Является ли язык $(\Sigma^* \setminus L) \cap c^* b^+ a^*$ КС-языком?

Решение.

4 $\langle$i$\rangle$.

Рассмотрим лемму о накачке для КС-языков.

Пусть данный язык – КС. Тогда для любого сколь угодно большого p будем рассматривать слово $w_p = a^p b^{p+1} c^{p+1}, |w_p| > p$

В данном случае $k = p, m = p + 1, n = p + 1$. Заметим, что $p < \min(p + 1, p + 1)$, т.е. $w_p \in L$

Рассмотрим все возможные разбиения слова w_p на $uvxyz$, $vy \neq \varepsilon$, $|vxy| \leqslant p$

1) $vxy = a^i, i \leqslant p$

   В этом случае $\exists k = 3 : uv^k xy^k z = a^{p+2i} b^{p+1} c^{p+1} \notin L$

2) $vxy = a^i a^j b^l, vx = a^i a^j, y = a^l b^m, i + j + l + m \leqslant p$

   В этом случае $\exists k = 3 : uv^k xy^k z = a^{p+2(i+l)} b^{p+1+2m} c^{p+1} \notin L$

3) $vxy = a^i a^j b^l b^m, v = a^i, x = a^j b^l, y = b^m, i + j + l + m \leqslant p$

   В этом случае $\exists k = 3 : uv^k xy^k z = a^{p+2i} b^{p+1+2m} c^{p+1} \notin L$

4) $vxy = a^i b^j b^k, v = a^i b^j, x = b^l, y = b^m, i + j + l + m \leqslant p$

   В этом случае $\exists k = 3 : uv^k xy^k z = a^{p+2i} b^{p+1+2(m+l)} c^{p+1} \notin L$

   Случаи (2),(3) и (4) – это случаи наложения vxy в разбиении на стыки последовательностей a, b, c, которые рассматриваются идентично при $vxy = b^i c^j$.

5) $vxy = b^i b^j b^l, i + j + l \leqslant p$

В этом случае $\exists k = 0 : uv^k xy^k z = a^p b^{p+1-i-l} c^{p+1} \notin L$

Случай наложения vxy целиком на последовательность c идентичен.

Получили, что для любого сколь угодно большого р существует слово, для которого для любого разбиения $uvxyz$ найдётся k, такое что $uv^k xy^k z \notin L$

Получили противоречие с леммой о накачке $\Rightarrow$ данный язык не КС.

По второму пункту заметим, что $\forall w \in L \to k > 0$. Отсюда получаем, что $\forall w' : w'[0] = c \to w' \in \Sigma^* \backslash L$

$\Rightarrow L' = (\Sigma^* \backslash L) \cap c^+(a|b)^* = c^+(a|b)^*$ Следовательно он регулярный (задан регулярным выражением $c^+(a|b)^*$), а значит и КС.

Критерии.

1.

- Разобраны только простые варианты разбиений (без пересечений): **1 балл**
- Полное решение, рассмотрены все разбиения, но есть ошибки: **2 балла**
- Полное решение, рассмотрены все разбиения: **4 балла**

2. Честно доказано, что пересечение L и регулярного языка R — это R и сказано, что регулярный язык является КС — **2 балла**. Без честного доказательства — **1 балл**.

Задача 5 (5). Пусть $\Sigma = \{\leftarrow, \rightarrow, \uparrow, \downarrow\}$. Определим $go : \Sigma^* \to \mathbb{Z}^2$ так:

$w \mapsto$ финальная точка маршрута по $\mathbb{Z}^2$, начатого в $(0,0)$ и пройденного вдоль пути, закодированного w

Например, $go(\leftarrow\rightarrow\uparrow\downarrow) = (0,0)$, а $go(\rightarrow\rightarrow\rightarrow\uparrow\uparrow\leftarrow) = (2,2)$.

Для $S \subset \mathbb{Z}^2$ определим $\mathsf{Ways}(S)$ как множество слов, кодирующих пути, стартующие в $(0,0)$ и не покидающие область S.

5 $\langle$i$\rangle$. Формально, обозначим $S = \mathsf{Ways}(\{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid y \geqslant x\})$. Постройте КС-грамматику или МП-автомат для языка

$$L_1 = \{w \in \{\rightarrow, \uparrow\}^* \mid go(w) = (x,x), x \geqslant 0\} \cap S.$$

5 $\langle$ii$\rangle$. Формально, обозначим $S = \mathsf{Ways}(\{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid y \leqslant x\})$. Постройте КС-грамматику или МП-автомат для языка

$$L_2 = \{w \in \{\leftarrow, \downarrow\}^* \mid go(w) = (x,x), x \leqslant 0\} \cap S.$$

5 $\langle$iii$\rangle$. Формально, обозначим $S_\geqslant = \mathsf{Ways}(\{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid y \geqslant -x\})$. Постройте КС-грамматику или МП-автомат для языка

$$L_3 = \{w \in \{\rightarrow, \downarrow\}^* \mid go(w) = (x,-x), x \geqslant 0\} \cap S.$$

5 $\langle$iv$\rangle$. Формально, обозначим $S_\leqslant = \mathsf{Ways}(\{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid y \leqslant -x\})$. Постройте КС-грамматику или МП-автомат для языка

$$L_4 = \{w \in \{\leftarrow, \uparrow\}^* \mid go(w) = (x,-x), x \leqslant 0\} \cap S.$$

Решение.

5 $\langle$i$\rangle$. Первый способ. Установим сначала некоторые свойства слов-путей из языка L_1. Для любого префикса p слова $w \in S$ справедливо $|p|_\uparrow \geqslant |p|_\rightarrow$, а для $w \in L_1$ справедливо $|w|_\uparrow = |w|_\rightarrow$, поскольку $go(w) = (x,x)$. Определим язык D_1 заменой в языке L_1 стрелок $\uparrow$ и $\rightarrow$ на скобки (и) соответсвенно. Упомянутые выше условия на стрелки в терминах скобок являются хорошо известным критерием правильных скобочных выражений, а потому язык D_1 порождается грамматикой, заданной правилами

$S \to (S)S \mid \varepsilon$, а L_1 соответсвенно порождается грамматикой G_1, заданной правилами $S \to\ \uparrow S \to S \mid \varepsilon$.

Отметим, что при повороте плоскости на 45° по часовой стрелке, мы получим, что пути из L_1 перейдут в пути по векторам $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ и $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ начинающиеся в точке $(0,0)$ и заканчивающиеся в точках на оси абсцисс и не опускающиеся ниже оси абсцисс. Эти пути известны как *пути Дика*, которые являются классической геометрической интерпретацией правильных скобочных выражений.

Второй способ. Во втором способе мы приведём явное доказательство того, что грамматика G_1 порождает язык L_1. $L(G_1) \subseteq L_1$. Докажем, что всякое слово длины N, выводящееся в G, лежит в L_1. Это можно доказать индукцией по N: пустое слово (длины 0) автоматически лежит в L_1, а слово длины N, выведенное грамматикой, имеет вид $\uparrow w_1 \to w_2$, где w_1, w_2 — выведенные грамматикой слова меньшей длины, то есть лежащие в L_1 по предложению индукции. Путь $\uparrow w_1 \to$ начинается и кончается на прямой $y = x$ и не опускается ниже её: на первом шаге путь поднимается на 1 вверх, не опускается ниже параллельной прямой $y = x + 1$, а затем возвращается обратно на прямую $y = x$. Тогда и $\uparrow w_1 \to w_2 \in L_1$.

$L_1 \subseteq L(G_1)$. Обозначим прямую $y = x$ через l. Отметим общее наблюдение: каждое слово из L_1 имеет вид $\uparrow u \to$. Действительно, если сначала пойти вправо, то путь пойдёт ниже прямой l, а если последний шаг «вверх», то путь придёт из-под прямой, что также невозможно.

Скажем, что собственный префикс p слова w ($p \neq \varepsilon$, $p \neq w$) является *элементарным*, если он оканчивается на точке прямой l и нет собственного префикса короче, оканчивающегося в некоторой точке l. Заметим, что элементарные префиксы лежат в L_1, а посему, если слово $w = ps$ и p — элементарный префикс, то $s \in L_1$ — пути начинающиеся и заканчивающиеся на прямой l остаются таковыми при замене стартовой точки на $(0,0)$.

Заметим также, что если слово $w \in L_1$ не имеет элементарного префикса, то $w =\uparrow u \to$ и $u \in L_1$: любой путь-префикс $\uparrow u$ не опускается ниже прямой $y = x + 1$ и заканчивается в точке этой прямой, а значит путь u ведёт в некоторую точку l и не опускается ниже l.

Итак, мы готовы доказать включение индукцией по длине слова $w \in L_1$. База: пустое слово принадлежит $L(G_1)$ по построению. Переход: пусть для всех слов из L короче n утверждение верно. Слово w длины n либо имеет элементарный префикс, либо нет. В первом случае $w = ps, p, s \in L_1$, $p =\uparrow u \to$ и в итоге $w =\uparrow u \to s$, слова p и s выводятся из G_1 по предложению индукции, а значит выводится и w: $S \Rightarrow\uparrow S \to S \overset{*}{\Rightarrow}\uparrow u \to s$. Если w не имеет элементарного префикса, то $w =\uparrow u \to, u \in L_1$ и u также выводится из G_1 по предложению индукции, а значит выводится и w (в предыдущем выводе берём $s = \varepsilon$).

Критерии.

Построение грамматики:

- **5 баллов** — верное решение
- **4 балла** — построена грамматика, доказано вложение $L_i \subset L(G)$
- **4 балла** — построена грамматика, доказано вложение $L(G) \subset L_i$, построен вывод элементарного слова
- **3 балла** — построена грамматика, доказано вложение $L(G) \subset L_i$
- **2 балла** — построена грамматика
- **0 баллов** — неверное решение

Построение автомата:

- **5 баллов** — верное решение
- **3.5 баллов** — автомат построен, доказано, что нужные слова принимаются
- **2 балла** — автомат построен, не доказана корректность
- **0 баллов** — неверное решение

Разумные соображения с попытками доказательства: $+$**1-2 очка** на усмотрение проверяющего. Особые случаи: в грамматике три или более стрелочки — **0 очков**.

1 Мини-задачи и вопросы

Общие критерии. В случае положительного ответа нужно привести доказательство, а в случае отрицательного – контрпример. Верное рассуждение без контрпримера оценивается в половину задачи (см. преамбулу). В задачах, оцениваемых в **3 балла**, третий балл ставится, если в доказательстве не было огрехов.

Задача 6(3). Обозначим через $\mathsf{Aut}(m)$ множество МП-автоматов, размер стека которых в процессе работы на любом слове не превосходит m.

6 $\langle$i$\rangle$. Существует ли МП-автомат $M \in \mathsf{Aut}(2017)$, распознающий язык $L = \{a^n b^n \mid n \geqslant 0\} \subseteq \{a, b\}^*$?

6 $\langle$ii$\rangle$. Существует ли МП-автомат $M \in \mathsf{Aut}(4034)$, распознающий язык $L = \{a^n b^{2n} a^n \mid n \geqslant 0\} \subseteq \{a, b\}^*$?

6 $\langle$iii$\rangle$. Существует ли МП-автомат $M \in \mathsf{Aut}(4034)$, распознающий язык $L = \{a^{2n} b^{2n} \mid n \geqslant 0\} \subseteq \{a, b\}^*$?

6 $\langle$iv$\rangle$. Существует ли МП-автомат $M \in \mathsf{Aut}(2017)$, распознающий язык $L = \{a^n b^{n+m} a^m \mid n, m \geqslant 0\} \subseteq \{a, b\}^*$?

Решение. В случае если стек ограничен, то его содержимое можно хранить в конечной памяти конечного автомата, поэтому для любого m язык, распознаваемый МП-автоматом $M \in \mathsf{Aut}(\mathsf{M})$ является регулярным. В свою очередь, язык L регулярным не является, как хорошо известно известно из курса. **Ответ:** «нет».

Приведём явную конструкцию конечного автомата. Поскольку размер стека МА M ограничен, конфигураций автомата, то есть в данном случае пар $(q, w) \in Q \times \Gamma^*$ конечное число. Поэтому таблица переходов МА M задает на них структуру конечного автомата. Формально говоря, пусть $M = (Q, \Sigma, \Gamma, \perp, Start, Final, T)$ — магазинный автомат, останавливающийся на пустом стеке, определим НКА $\widetilde{M}$, принимающий $L(M)$:

- $\widehat{Q} = Q \times (\{\varepsilon\} \cup \perp (\varepsilon + \Gamma)^m)$ — множество состояний НКА, конфигурации M,
- Σ — алфавит, данный в условии,
- $\widehat{Start} = (Start, \perp)$ — начальная конфигурация;

- $\widehat{Final} = Q \times \{\varepsilon\}$ — финальные состояния суть конфигурации, в которых оказывается M принимает слово;
- $\widehat{T}$ — таблица переходов, устроенная так: пусть старая таблица переходов $T : (q_1, x, X) \to (q_2, \gamma)$ определяет функцию

  $$f : (\Gamma\cup\{\varepsilon\}\cup\{\perp\})\times(\{\varepsilon\}\cup\perp(\varepsilon+\Gamma)^m) \to (\{\varepsilon\}\cup\perp(\varepsilon+\Gamma)^m),\ (X, wX) \mapsto w\gamma,$$

  тогда зададим

  $$\forall q \in Q, w \in (\varepsilon+\perp)(\varepsilon+\Gamma)^{m-2}, u \in \Gamma\ \widehat{T}((q, uw), x) = \{(T(q, u, x), f(u, uw))\}$$

Осталось проверить корректность конструкции. То, что определен НКА, очевидно, надо лишь проверить, что его язык есть в точности $L(M)$. Индукцией по $|w|$ можно показать, что

$$\exists q \in Q, u \in \Gamma^*\ (Start, \perp) \xrightarrow{w}_M (q, u) \iff (Start, \perp) \xrightarrow{w}_{\widetilde{M}} (q, u)$$

База этого утверждения очевидна; далее, слово длины $|w|+1$ получается дописыванием к w некоторой буквы $\alpha \in \Sigma$, тогда

$$(Start, \perp) \xrightarrow{w}_M (q, u) \xrightarrow{\alpha}_M T(q, \alpha, u[0]) \iff (Start, \perp) \xrightarrow{w}_{\widetilde{Aut}} (q, u) \xrightarrow{\alpha}_{\widetilde{M}} \widehat{T}((q, u), \alpha)$$

и $\widehat{T}((q, u), \alpha) = (T(q, \alpha, u), f(u[0], u)) = T(q, \alpha, u[0])$. Тогда $w \in L(M) \iff w \in L(\widetilde{M})$. Таким образом, $L(M)$ есть язык некоторого НКА, а значит, регулярен.

Критерии.

- **3 балла** — верное решение
- **+ 2 балла** — построен конечный автомат
- **+ 1 балл** — доказана корректность построенного автомата
- **0 баллов** — неверное решение

Достаточно сослаться на нерегулярность языка как на факт из курса.

Особые случаи — анимизм автоматов «автоматы знают, автоматы думают» — **0 очков**.

Задача 7 (3). Пусть $\Sigma = \{a, b\}$.

7 $\langle$i$\rangle$. Существуют ли два разных языка $X \subseteq \Sigma^*$, таких что $X = aXb + \varepsilon$.

7 $\langle$ii$\rangle$. Существуют ли два разных языка X таких, что $X = aXb$?

7 $\langle$iii$\rangle$. Существуют ли два разных языка X таких, что $X = aXb + bXa$?

7 $\langle$iv$\rangle$. Существуют ли два разных языка X таких, что $X = aXXb$?

Решение.

Во всех вариантах ответ отрицательный.

Первый вариант: если бы существовали два таких языка A и B, что $A = aAb + \varepsilon$ и $B = aBb + \varepsilon$, то в обоих языках нет слов нечетной длины и ровно одно слово четной длины $2n$, а именно $a^n b^n$. Ведь если $X = aXb + \varepsilon$, то $X \cap \Sigma^{2n} = a(X \cap \Sigma^{2n-2})b$, а $X \cap \Sigma^0 = \varepsilon$, поэтому если $X \cap \Sigma^{2n-2} = a^{n-1}b^{n-1}$, то $X \cap \Sigma^{2n} = a^n b^n$. Поэтому $A = B$, двух разных решений уравнения не существует.

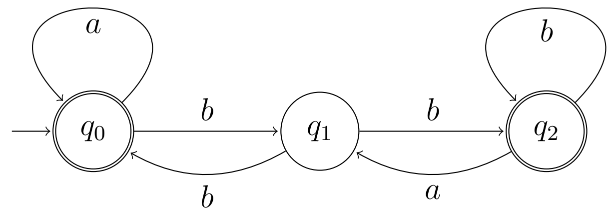
В остальных вариантах единственное решение — пустой язык. Действительно, если бы существовал непустой язык L, удовлетворявший данному уравнению (например, $X = aXb$, как во втором варианте), то существует слово $w \in L$ минимальной длины; тогда оно же есть слово вида asb, где $s \in L$, следовательно, $|s| = |w| - 2$ и w имеет не наименьшую длину среди слов языка L.

Критерии.

- **3 балла** — верное решение, не использующее КС грамматик (по индукции показано, что ровно определенные слова должны быть)
- **2 балла** — доказывается, что язык не содержит слова минимальной длины, после чего заключается, что такого языка не бывает.
- **+1 балл** — доказано, что данный язык удовлетворяет данному уравнению
- **+1 балл** — построена КС грамматика соответствующего языка
- **+1 балл** — сказано, что язык единственный и он пустой (но не вносит вклад +1 к другому частичному решению).
- **0 баллов** — неверное решение

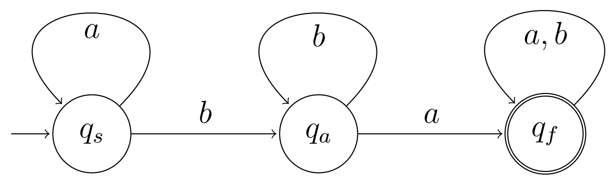
Задача 8 (3). Автомат $\mathcal{A}$ задан диаграммой:

8 ⟨i⟩.



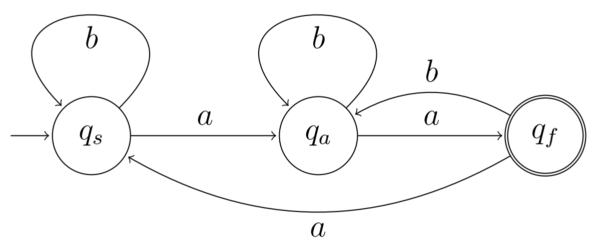
Верно ли, что язык $L(\mathcal{A})$ не порождается ни одной LL(1)-грамматикой?

8 ⟨ii⟩.



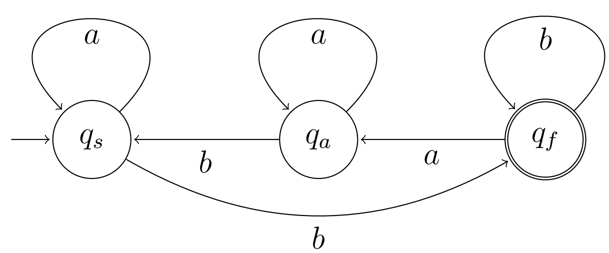
Верно ли, что язык $L(\mathcal{A})$ не порождается ни одной LL(1)-грамматикой?

8 ⟨iii⟩.



Верно ли, что язык $L(\mathcal{A})$ не порождается ни одной LL(1)-грамматикой?

8 ⟨iv⟩.



Верно ли, что язык $L(\mathcal{A})$ не порождается ни одной LL(1)-грамматикой?

Решение.

Регулярный язык является языком некоторой праволинейной грамматики, построенной по ДКА. Правила этой грамматики имеют вид $Q_i \to xQ_j$ для $x \in \Sigma$ и соответствуют переходам $q_i \xrightarrow{x} q_j$ в ДКА. В силу детерминированности по Q_i и x нетерминал Q_j определяется однозначно, поэтому по первому символу непрочитанной части следующее правило угадывается однозначно.

Проведем наши рассуждения более формально. Мы рассматриваем праволинейную грамматику $(\{Q_i\}, \Sigma, \{Q_i \to xQ_j\}, Q_0)$, построенную по полному ДКА $Q = (\{q_i\}, \Sigma, q_0, Final, T)$, в ней правила $Q_i \to xQ_j$ соответствуют правилам $q_i \xrightarrow{x} q_j$. Все выведенные строчки в такой грамматике имеют вид wQ_i, где $w \in \Sigma^*$, а $q_0 \xrightarrow{w} q_i$ в данном ДКА, то есть q_i — состояние, в котором оказывается ДКА при прочтении слова w. [Последнее можно доказать по индукции: база очевидна, а если утверждение верно для всех слов длины n, то $n+1$-ый шаг вывода имеет вид $wQ_i \to wxQ_j$, где $q_i \xrightarrow{x} q_j$.] Тогда для любых двух левых выводов $S \Rightarrow^* wQ_i \Rightarrow wxQ_j \Rightarrow^* ww_1$ и $S \Rightarrow^* wQ_i \Rightarrow wxQ_k \Rightarrow^* ww_2$ из $w_1[1] = w_2[1] = x$ следует $Q_j = Q_k$, то есть эта праволинейная грамматика является LL(1)-грамматикой.

Критерии. Решение, указанное здесь:

- **3 балла** — верное решение
- **1.5 балла** — сказано, что для любого регулярного языка праволинейная грамматика LL(1), но не доказано
- **0 баллов** — неверное решение

Проверка LL(1)-условия:

- **3 балла** — верное решение
- **-0.5 балла** — за каждую ошибку при проверке LL(1)-условия
- **1 балла** — построена грамматика, не проверено LL(1)-условие
- **0 баллов** — неверное решение

Задача 9 (2).

9 ⟨i⟩. Существует ли магазинный автомат с не более чем 100 состояниями, распознающий язык $L = \{a^{100n}b^{100n} \mid n \geqslant 0\}$.

9 ⟨ii⟩. Существует ли магазинный автомат с не более чем 200 состояниями, распознающий язык $\{a^{200n}b^{400n}a^{200n}|n \in \mathbb{N}\}$?

9 ⟨iii⟩. Существует ли магазинный автомат с не более чем 50 состояниями, распознающий язык $\{a^{1000n}b^{1000n}|n \in \mathbb{N}\}$?

9 ⟨iv⟩. Существует ли магазинный автомат с не более чем 150 состояниями, распознающий язык $\{a^{300n}b^{800n}a^{500n}|n \in \mathbb{N}\}$?

Решение.

В первом и третьем варианте ответ «да», так как языки $\{a^{100n}b^{100n}|n \in \mathbb{N}\}$ и $\{a^{1000n}b^{1000n}|n \in \mathbb{N}\}$ контекстно свободные (они порождаются грамматиками $\{\{S\},\ \{a, b\},\ \{S \to a^{100}Sb^{100} \mid \varepsilon,\ S\},\ S\}$ и $\{\{S\},\ \{a, b\},\ \{S \to \to a^{1000}Sb^{1000} \mid \varepsilon,\ S\},\ S\})$, таким образом, они распознаются некоторым МП автоматом с одним состоянием.

Во втором и четвертом варианте ответ «нет»: языки $\{a^{200n}b^{400n}a^{200n} \mid n \in \mathbb{N}\}$ и $\{a^{300n}b^{800n}a^{500n} \mid n \in \mathbb{N}\}$ не контекстно свободны. Действительно, пусть существует $p \in \mathbb{N}$ такое, что любое слово большей длины допускает разбиение $xyzvw$ такое, что $|yzv| < p$, $yv \neq \varepsilon$ и $xy^izv^iw \in \{a^{200n}b^{400n}a^{200n}|n \in \mathbb{N}\}$ для всех $i \in \mathbb{N}$. Тогда нашлось бы слово $a^{200p}b^{400p}a^{200p}$ (в четвертом варианте — $a^{300p}b^{800p}a^{500p}$), тогда либо $yv = a^k$, либо $yv = b^k$, либо $y = a^k$ и $v = b^l$, либо $y = b^k$ и $v = a^l$, в любом из этих случаев получится, что $a^xb^ya^z \in \{a^{200n}b^{400n}a^{200n}|n \in \mathbb{N}\}$ для x, y, z таких, что либо $x \neq z$, либо $2x \neq y$, либо $y \neq 2z$ (слово $|yzv|$ по длине меньше $400p$). Если же либо y, либо v равен либо a^ib^j, либо b^ia^j (для $i, j \neq 0$), то получится некоторое $xy^2zv^2w \in a^+b^+a^+b^+a^+$, которое не может лежать в $\{a^nb^na^n|n \in \mathbb{N}\}$. (Вопрос из четвертого варианта разбирается аналогично.)

Критерии.

Если ответ «да» (i,iii):

- **2 балла** — верное решение
- **1 балл** — сказано, что МП с одним состоянием есть, нет доказательства, что данный язык КС

- **0 баллов** — неверное решение

Если ответ «нет» (ii,iv):

- **2 балла** — верное решение
- **1 балл** — неполный разбор случаев разбиений в лемме о накачке
- **0 баллов** — неверное решение