Задание 4

Коновалов Андрей, 074

1	2	3	4	5	6	σ

Задача 1

(ii). Контрпример: язык L в алфавите $\{a, b\}$ задается регулярным выражением: $a \cap b$, морфизм h определяется: $h(a) = a$, $h(b) = a$.

Видно, что язык $L = \varnothing$, а значит $h(L) = \varnothing$. Подставим в регулярное выражение значения морфизма h на литералах, получим: $h(a) \cap h(b) = a \cap a = a$. Но $\{a\} \neq \varnothing$.

Задача 2

(i). Пересечем данный язык L_1 с регулярным языком, заданным регулярным выражением $0(1)^*$. Получим язык $L = \{01^{2^i} | i \geq 0\}$.

Докажем, что язык L - нерегулярный. Допустим, что он регулярный, тогда он удовлетворяет лемме о разрастании с некоторой константой p. Возьмем слово $w = 01^{2^p}$. Его длина больше, чем p, следовательно существует такое разбиение $w = xyz$, что $0 < |y| \leq C$ и для всех $i \geq 0$, $xy^iz \in L$. Возможны только следующие случаи:

1. $x = \varepsilon$. В этом случае y имеет префикс 0, поскольку $|y| > 0$. Но тогда при $i = 2$ слово xy^iz будет содержать две буквы 0, следовательно $\notin L$.

2. $x \neq \varepsilon$. В этом случае $y = 1^n$ для некоторого $n \geq 1$. Заметим, что количество слов вида xy^iz, длиной не больше, чем некотрое число C растет, в зависимости от C, линейно. При этом число слов в языке L, длиной не более C растет логарифмически. Следовательно, язык L не может содержать все слова вида xy^iz.

В обоих случаях мы пришли к противоречию, следовательно L - нерегулярный, а значит и язык L_1 - нерегулярный.

(ii). Пересечем данный язык L_2 с регулярным языком, заданным регулярным выражением $0(1)^*0(1)^*$. Поссмотрим на результирующий язык L. Поскольку любое его слово w должно быть квадратом некоторого слова v, и w должно содержать ровно две буквы 0 можно сделать вывод о том, что v содерджит ровно одну букву 0. Поскольку первая половина w должна начинаться с буквы 0, то и вторая тоже. А значит w имеет вид $0(1)^n0(1)^n$. Получаем $L = \{0(1)^n0(1)^n | n \geq 0\}$.

Докажем, что язык L - нерегулярный. Допустим, что он регулярный, тогда он удовлетворяет лемме о разрастании с некоторой константой p. Возьмем слово $w = 01^p01^p$. Его длина больше, чем p, следовательно существует такое разбиение $w = xyz$, что $0 < |y| \leq C$ и для всех $i \geq 0$, $xy^iz \in L$. Возможны только следующие случаи:

1. y содержит букву 0. В этом случае при $i = 2$ поличество букв 0 в слове xy^iz будет больше двух, а значит $xy^2z \notin L$.

2. $y = 1^n$ для некоторого $n \geq 1$. Но в этом случае для $i = 2$ количество букв 1 до второго вхожения буквы 0 будет отличаться от количества букв 1 после этого вхождения на n, а значит $xy^2z \notin L$.

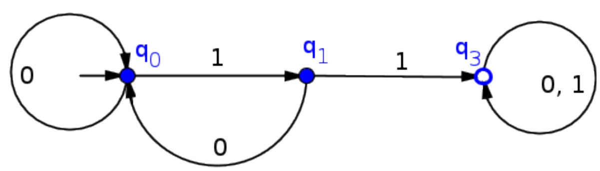
В обоих случаях мы пришли к противоречию, следовательно L - нерегулярный, а значит и язык L_2 - нерегулярный.

Задача 3

Нет, поскольку при $n = m = 0 \rightarrow nm = 0$, а значит $(0, 0) \notin R$, следовательно R - не рефлексивное, следовательно R - не отношение эквивалентности.

Задача 5

(i) Построим ПДКА A для L. Он изображен на диаграмме ниже.



Докажем его корректность по индукции по длине n слова w.

База. При $n = 0$ автомат принимает слово ε. При $n = 1$ автомат принимает слова 0 и 1. Свойство выполняется. База доказана.

Переход. Пусть автомат принимает только те слова длины меньше n, которые не содержат двух букв 1 подряд. Докажем, что аналогичное утверждение выполняется для слов длины n.

Возьмем слово w длины n. Пропустим через A префикс v длины $n - 1$ слова w. Возможны два случая:

1. A не принимает v. Это означает, что v содержит две буквы 1 подряд, а значит и w их содержит и не должно быть принято. Заметим, что A после обработки v окажется в единственном нефинальном состоянии q_3. Какой бы ни была последняя буква слова w, A так и останется в q_2, а значит w не будет принято. Необходимое свойство выполняется.

2. A принимает v, а значит находится либо в состоянии q_0 или в состояниее q_2. Если последняя буква v была 1, а последняя буква w тоже 1, то A перейдет в q_3 и w не будет принято. Если последняя буква v была 1, а последней буквой w является 0, то A перейдет в q_0 и w будет принято. Если последняя буква слова v не существовала или была 0, то A после обработки v окажется в q_0, и при переходе по последней букве слова w перейдет или в q_0 или в q_1 и w будет принятно. Необходимое свойство выполняется.

Переход доказан.

(ii) Построим праволинейную грамматику $G = (\{Q_0, Q_1, Q_2\}, \{0, 1, \varepsilon\}, P, Q_0)$ по автомату A. Множество выводов P будет выглядеть следующим образом:

$$
\begin{aligned}
&Q_0 \to 0Q_0\\
&Q_0 \to 1Q_1\\
&Q_1 \to 0Q_0\\
&Q_1 \to 1Q_2\\
&Q_2 \to 0Q_2\\
&Q_2 \to 1Q_2\\
&Q_0 \to \varepsilon\\
&Q_1 \to \varepsilon
\end{aligned}
$$

Составим систему регулярных уравнений по полученной грамматике.

$$
\begin{cases}
Q_0 = 0Q_0 + 1Q_1 + \varepsilon\\
Q_1 = 0Q_0 + 1Q_2 + \varepsilon\\
Q_2 = 0Q_2 + 1Q_2
\end{cases}
$$

Найдем наименьшую неподвижную точку этой системы.

$$
\begin{aligned}
&Q_2 = (0+1)Q_2\\
\varepsilon \notin \{0,1\} \Rightarrow\ &Q_2 = \varnothing\\
&Q_1 = 0Q_0 + \varepsilon\\
&Q_0 = 0Q_0 + 1Q_1 + \varepsilon\\
&Q_0 = 0Q_0 + 1(0Q_0 + \varepsilon) + \varepsilon\\
&Q_0 = (0+10)Q_0 + (1+\varepsilon)\\
&Q_0 = (0+10)^*(1+\varepsilon)\\
&Q_1 = 0(0+10)^*(1+\varepsilon) + \varepsilon
\end{aligned}
$$

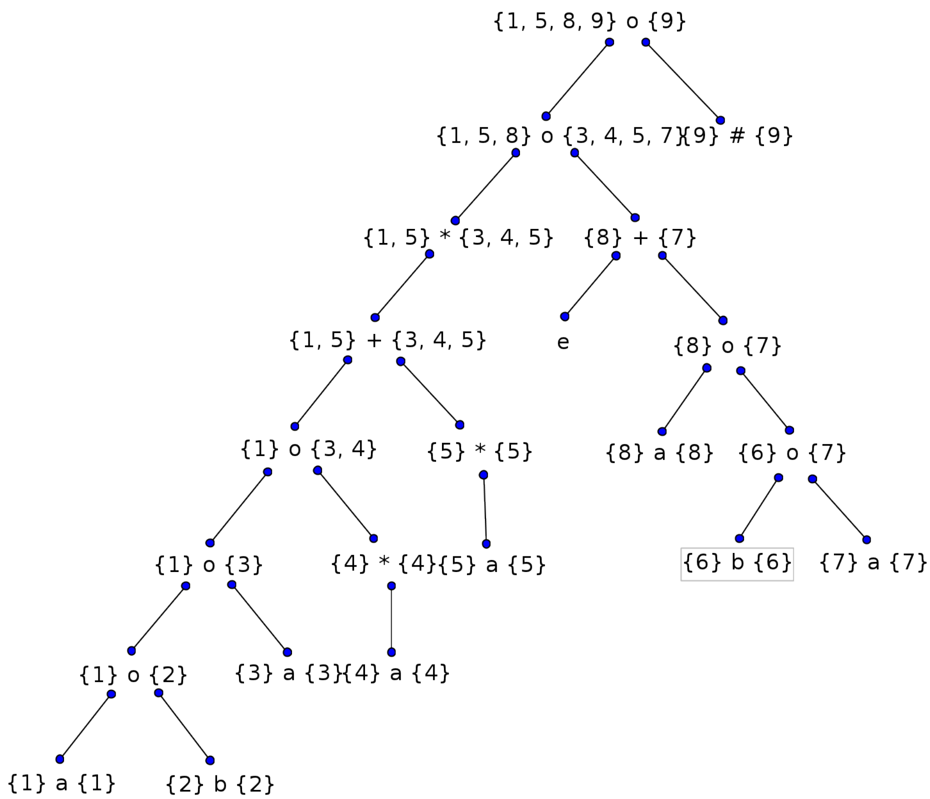
Наименьшая неподвижная точка:

$$
\begin{cases}
Q_0 = (0+10)^*(1+\varepsilon)\\
Q_1 = 0(0+10)^*(1+\varepsilon) + \varepsilon\\
Q_2 = \varnothing
\end{cases}
$$

(iii) Подставим в регулярное выражение $Q_0 = (0+10)^*(1+\varepsilon)$ вместо литералов регулярные выражения для языков L_0 и L_1 соответственно: $0 \to a^*$, $1 \to aba$. Получим регулярное выражение $R = (a^* + aba(a)^*)^*(aba+\varepsilon)$.

Воспользуемся алгоритмом 3.3.3 из книги Серебрякова для построения автомата по R. Сначала дополним регулярное выражение символом $\#$, получим $(a^* + aba(a)^*)^*(aba+\varepsilon)\#$. Теперь построим синтаксическое дерево по полученному регулярному выражению.

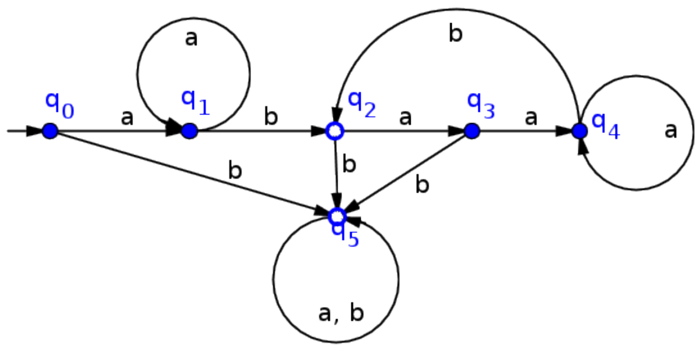
На диаграмме ниже изображено построенное синтаксическое дерево с результатом вычисления функций $firstpos$ и $lastpos$, значение которых записаны соответственно слева и справа от каждого узла дерева.



Теперь вычислим значения функции $followpos$. Результаты записаны в следующую таблицу:

позиция	followpos
1	$\{2\}$
2	$\{3\}$
3	$\{1, 4, 5, 8, 9\}$
4	$\{1, 4, 5, 8, 9\}$
5	$\{1, 5, 8, 9\}$
6	$\{7\}$
7	$\{9\}$
8	$\{6\}$
9	$\varnothing$

По полученным значениям $followpos$ построим автомат.



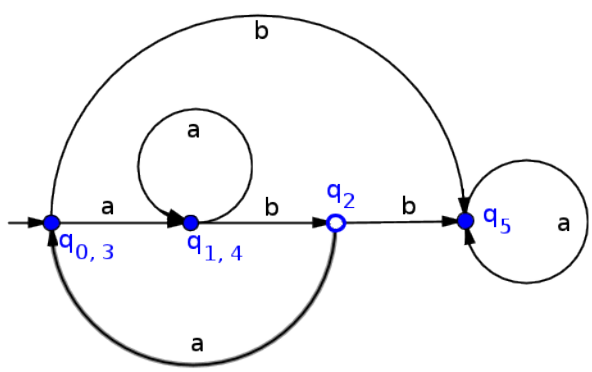
На диаграмме выше, изображен автомат, уже дополненный до полного состоянием q_5. При этом каждому из состояний q_0, q_1, q_2, q_3, q_4 соответствует множество позиций синтаксического дерева, указанное в следующей таблице:

состояние	множество позиций
q_0	$\{1,5,8,9\}$
q_1	$\{1,2,5,6,8,9\}$
q_2	$\{3,7\}$
q_3	$\{1,2,4,5,6,8,9\}$
q_4	$\{1,4,5,8,9\}$

(iv) По алгоритму, описанному в теории построим минимальный эквивалентный автомат $min(A)$. В исходном автомате нет недостижимых состояний, в чем легко убедиться. Теперь построим индуктивное отношение эквивалентности $R \equiv^{|Q|-2}$, где $|Q|-2 = 6-2 = 4$.

$$q_0 \equiv^0 q_1 \equiv^0 q_3 \equiv^0 q_4; \;\; q_2 \equiv^0 q_5;$$
$$q_0 \equiv^1 q_1 \equiv^1 q_3 \equiv^1 q_4; \;\; q_2; \;\; q_5;$$
$$q_0 \equiv^2 q_3; \;\; q_1 \equiv^2 q_4; \;\; q_2; \;\; q_5;$$
$$q_0 \equiv^3 q_3; \;\; q_1 \equiv^3 q_4; \;\; q_2; \;\; q_5;$$
$$q_0 \equiv^4 q_3; \;\; q_1 \equiv^4 q_4; \;\; q_2; \;\; q_5;$$

Теперь объединим эквивалентные состояния и построим $min(A)$.



(v) Для каждого состояния $min(A)$ построим достигающие цепочки.

состояние	достигающая цепочка
$q_{0,3}$	ε
$q_{1,4}$	a
q_2	ab
q_5	abb

Для каждой пары различных состояний построим различающие цепочки.

	$q_{0,3}$	$q_{1,4}$	q_2	q_5
$q_{0,3}$	X	ba	ε	ε
$q_{1,4}$	ba	X	ε	ε
q_2	ε	ε	X	a
q_5	ε	ε	a	X