

МИНОБРНАУКИ РОССИИ
Федеральное государственное автономное образовательное учреждение высшего
профессионального образования «Южный федеральный университет»

Институт математики, механики и компьютерных наук им. И.И. Воровича

Кафедра алгебры и дискретной математики

КУРСОВАЯ РАБОТА

по предмету: «Теория автоматов и формальных языков»

Выполнил:

Студент 3 курса, 9 группы

Стребежев Игорь

Проверил:

к.т.н., ст. преп. кафедры АДМ

Е.В. Алымова

Ростов-на-Дону
2015

Задание 1

Вариант 18. Язык над алфавитом $\Sigma = \{0, 1\}$, состоящий из всех слов, в которых после третьей слева единицы стоит четное число групп вида 01.

- (i) Построить ПЛ-грамматику G , порождающую L .

$$\begin{aligned} S &\rightarrow 0S \mid 1A, \\ A &\rightarrow 0A \mid 1B, \\ B &\rightarrow 0B \mid 1C \mid 1, \\ C &\rightarrow 0101C \mid 0101; \end{aligned} \tag{1}$$

- (ii) Доказать вложения $L \subseteq L(G)$, $L(G) \subseteq L$.

$L \subseteq L(G)$: Рассмотрим дерево выводов грамматики.

$S \Rightarrow 0S \Rightarrow 00S \Rightarrow 000S \Rightarrow \dots$ — нули разрешены в начале слов в L .

$S \Rightarrow 1A \Rightarrow 10A \Rightarrow 100A \Rightarrow \dots$ — генерация первой единицы и последовательности нулей.

Первые три правила грамматики одинаковы и образуют три пары из последовательности (возможно пустой) нулей и единицы. Что также удовлетворяет описанию языка L .

$11B \Rightarrow 111$ — в таком случае количество групп вида 01 будет равно нулю (чётно), принадлежит языку.

$11B \Rightarrow 111C \Rightarrow 111\ 0101$ — генерация пары групп.

$11B \Rightarrow 111C \Rightarrow 111\ 0101C \Rightarrow 111\ 0101\ 0101$ — всегда чётное количество групп.

Таким образом все слова, порождаемые грамматикой G принадлежат языку L , то есть $L \subseteq L(G)$.

$L(G) \subseteq L$: Язык L имеет структуру $(0^*1)(0^*1)(0^*1)(010)^*$, каждая цепочка этого языка описывает все терминальные выводы из правил грамматики, как показано в предыдущем пункте, что и доказывает вложение.

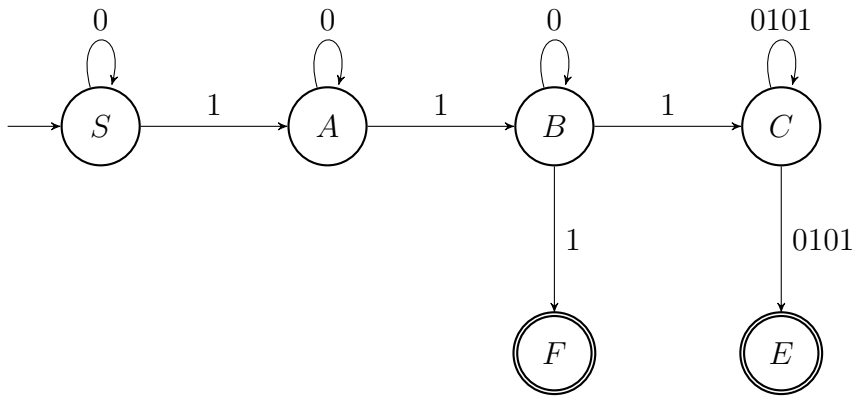
- (iii) Путём решения системы линейных уравнений с регулярными коэффициентами построить регулярное выражение, описывающее L .

$$\begin{aligned} S &= 0S + 1A, \\ A &= 0A + 1B, \\ B &= 0B + 1C + 1, \\ C &= 0101C + 0101; \end{aligned}$$

$$\begin{aligned} C &= (0101)^*0101 = 0101^+, \\ B &= 0B + 1(0101^+ + \varepsilon) = 0^*1(0101^+ + \varepsilon), \\ A &= 0A + 1(0^*1(0101^+ + \varepsilon)) = 0^*1(0^*1(0101^+ + \varepsilon)), \\ S &= 0S + 1(0^*1(0^*1(0101^+ + \varepsilon))) = 0^*1(0^*1(0^*1(0101^+ + \varepsilon))); \end{aligned} \tag{2}$$

$$0^*1(0^*1(0^*1(0101^+ + \varepsilon))) = (0^*1)(0^*1)(0^*1)(0101)^*.$$

(iv) Построить НКА или e -НКА M^{ND} , распознающий язык L , предъявить его граф.



(v) Построить ДКА M^D путём детерминизации M^{ND} , предъявить его граф.

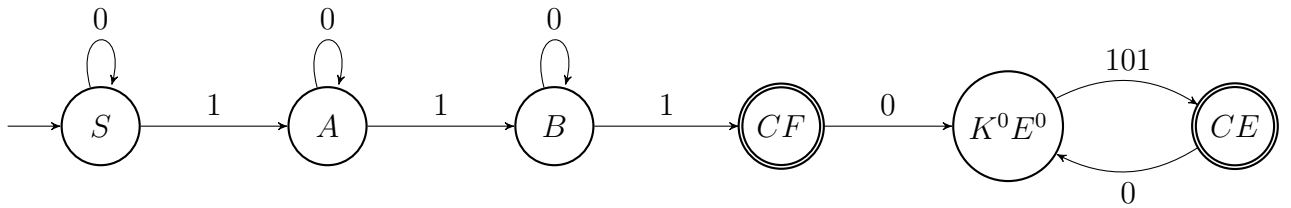
Построим таблицу переходов M^{ND} .

$\delta' \mid$	S	A	B	C	K^0	K^{01}	K^{010}	E^0	E^{01}	E^{010}	E	F
0 \mid	S	A	B	$\{K^0, E^0\}$	—	K^{010}	—	—	E^{010}	—	—	—
1 \mid	A	B	$\{C, F\}$	—	K^{01}	—	C	E^{01}	—	E	—	—

Расширим таблицу при приведении автомата в M^D .

$\delta' \mid$	S	A	B	$\{C, F\}$	$\{K^0, E^0\}$	$\{K^{01}, E^{01}\}$	$\{K^{010}, E^{010}\}$	$\{C, E\}$
0 \mid	S	A	B	$\{K^0, E^0\}$	—	$\{K^{010}, E^{010}\}$	—	$\{K^0, E^0\}$
1 \mid	A	B	$\{C, F\}$	—	$\{K^{01}, E^{01}\}$	—	$\{C, E\}$	—

Отобразим этот автомат на диаграмме.

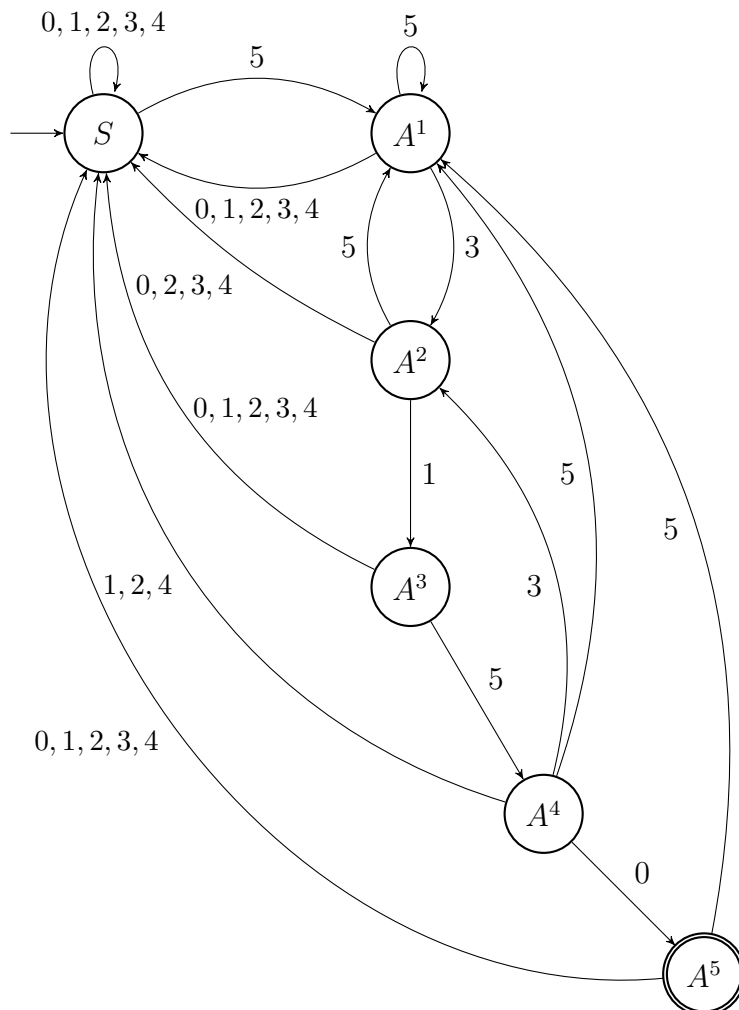
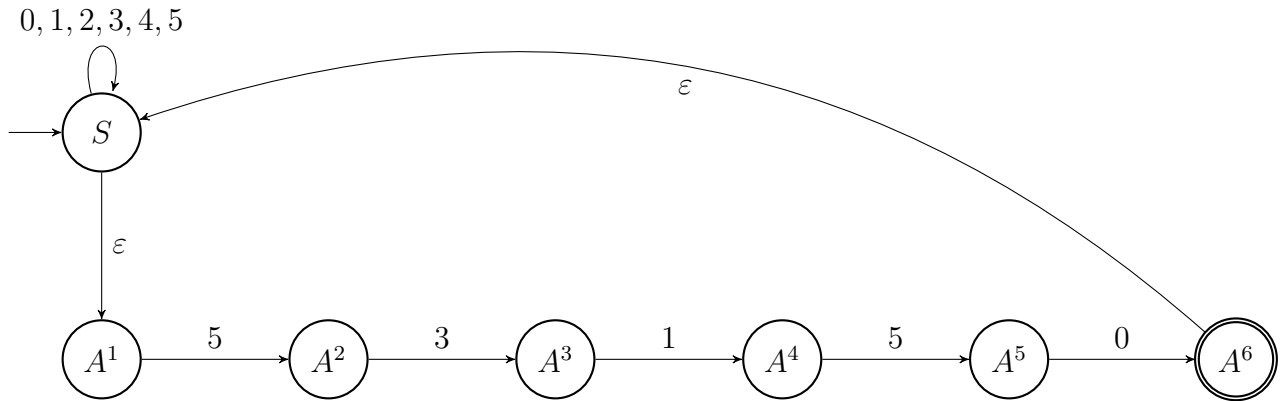


Задание 2

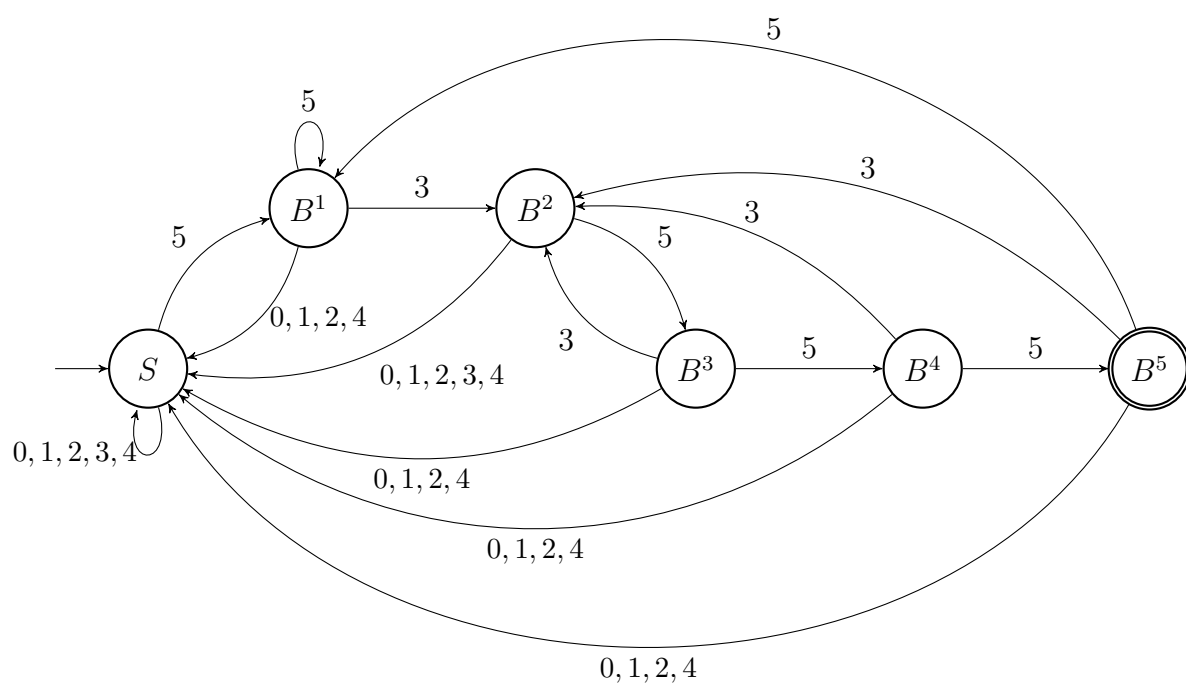
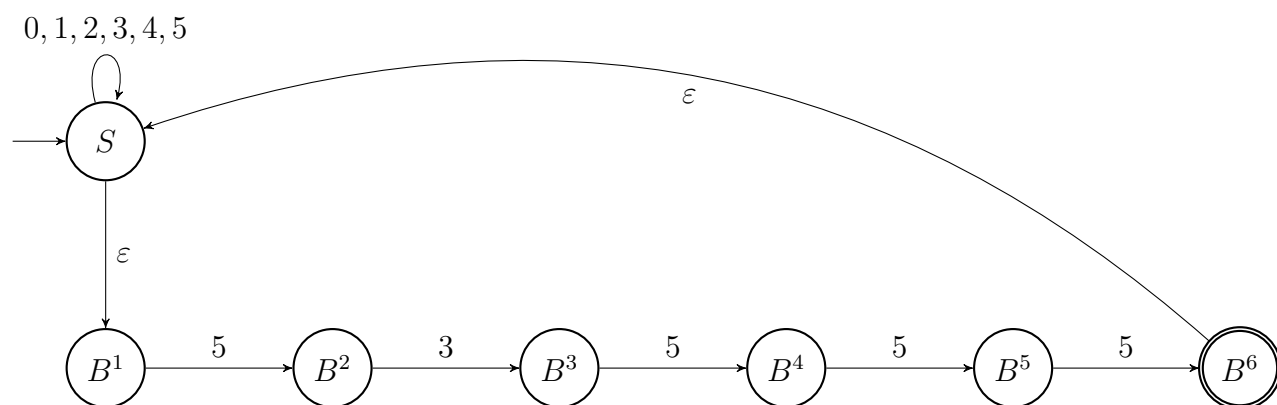
Вариант 29. $\Sigma = \{0, 1, 2, 3, 4, 5\}$, $A = \{53150, 53555, 5510, 0001\}$.

1. Для каждого слова $w_i \in A$ построить НКА M_i^{ND} , распознающий наличие в произвольной строке $s \in \Sigma^*$ подстроки w_i .
2. Для каждого НКА M_i^{ND} построить соответствующий ДКА M_i^D .

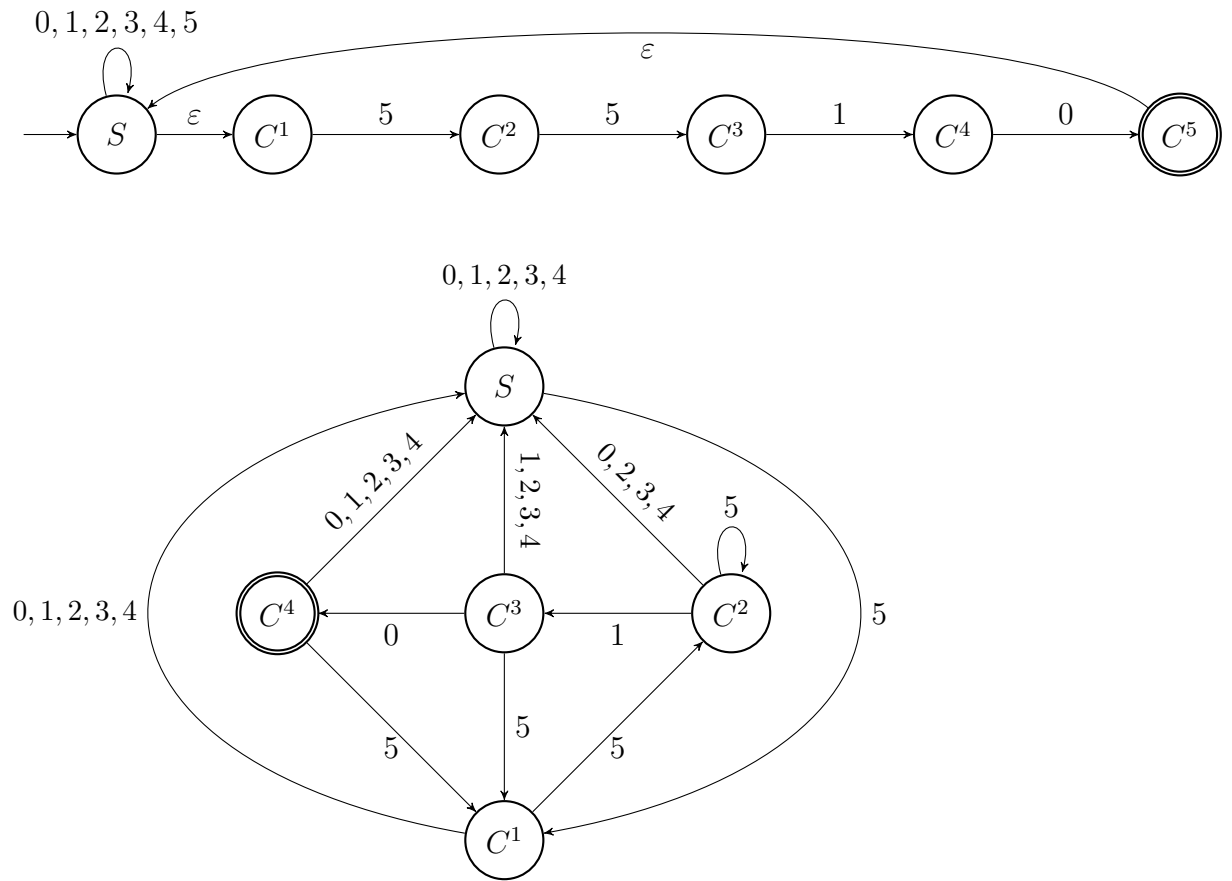
Слово 53150.



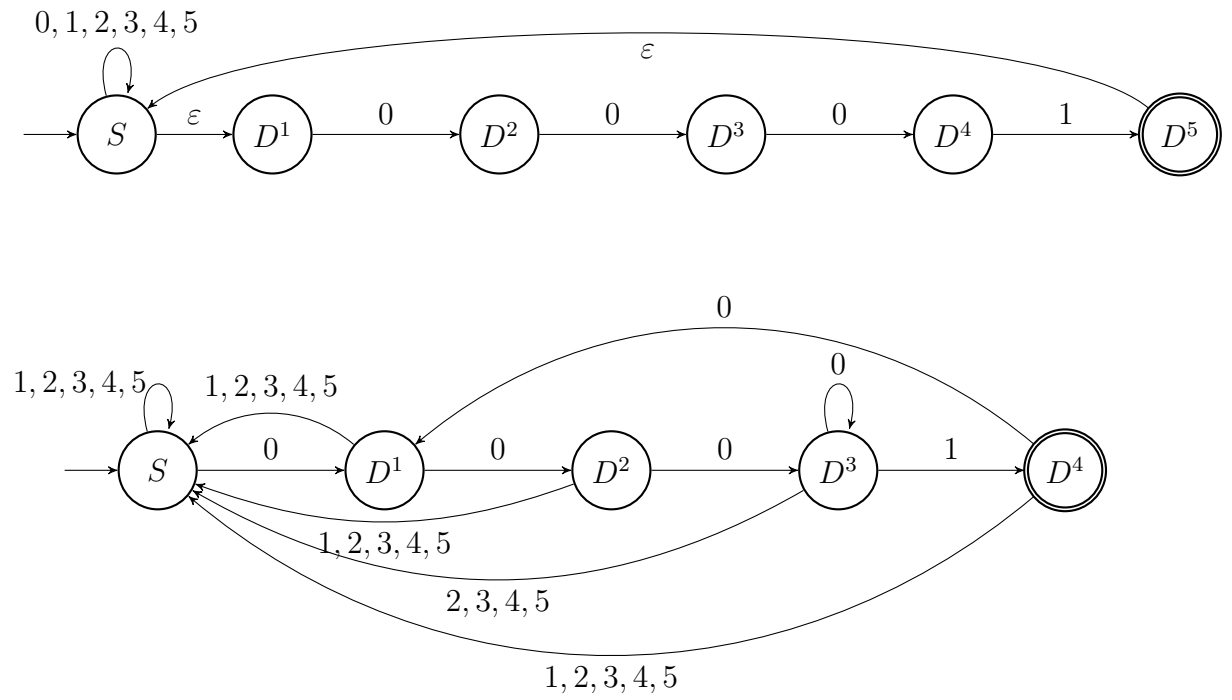
Слово 53555.



Слово 5510.



Слово 0001.



Задание 3

$$L_1 = b^*bba^*(ab + b)^*, L_2 = a^*ba^*ab.$$

- (i) Вычислить регулярное выражение, определяющее язык $L_1 \cap L_2$.

Рассмотрим множество слов из L_2 таких, что их первая буква равна a , $M = \{w \mid w \in L_2, w^0 = a\}$. Очевидно, что $L_1 \cap M = \emptyset$, так как слова из L_1 начинаются с буквы b . Уберём это множество из рассмотрения, пусть $L'_2 = ba^*ab$, где $L'_2 \cup M = L_2$.

В множестве L'_2 все слова начинаются со строки "ba", а в множестве L_1 — "bb". Из этого следует, что L_1 и L'_2 не пересекаются, а значит $L_1 \cap L_2 = \emptyset$.

- (ii) Вычислить регулярное выражение, определяющее язык $L_1 \triangle L_2$.

Согласно определению симметрической разности имеем:

$$L_1 \triangle L_2 = (L_1 \cup L_2) \setminus (L_1 \cap L_2) = L_1 \cup L_2 = b^*bba^*(ab + b)^* + a^*ba^*ab.$$

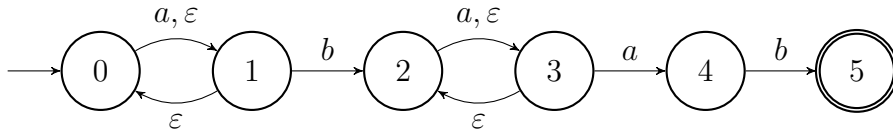
- (iii) Определить: совпадают ли языки L_1 и L_2 , является ли L_1 дополнением L_2 .

Так как $L_1 \cap L_2 = \emptyset$, то $L_1 \neq L_2$.

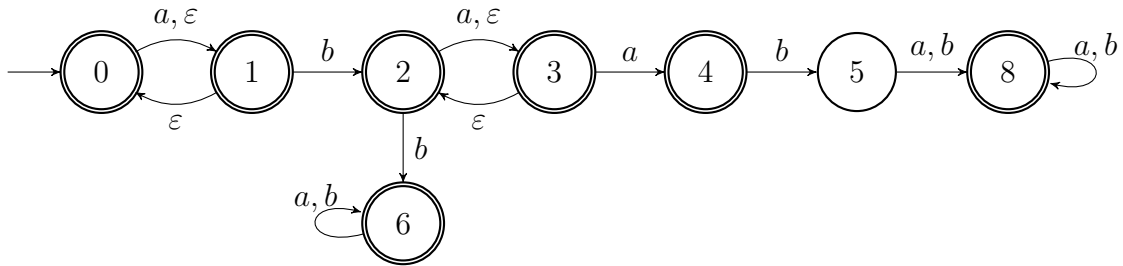
Предположим, L_1 является дополнением L_2 . Рассмотрим слово $w = "ba"$. Видно, что $w \notin L_1$, $w \notin L_2 \Rightarrow$ наше предположение неверно.

- (iv) Построить e -НКА, распознающий один из языков L_1^R или L_2^R .

Построим L_2 :

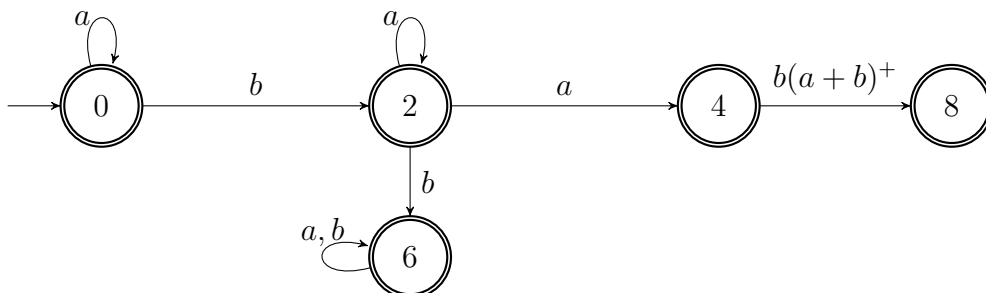


Построим L_2^R :



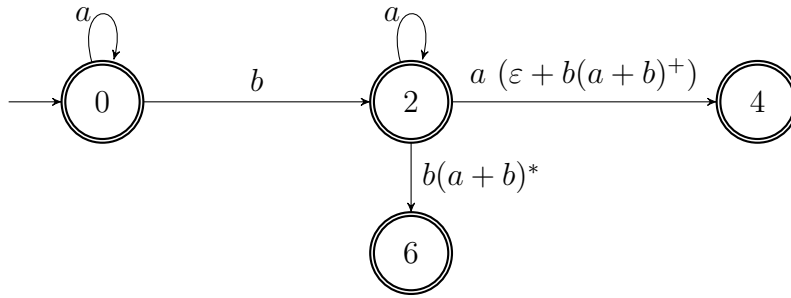
- (v) Вычислить регулярное выражение по построенному e -НКА.

Упростим построенный в предыдущем пункте автомат, используя метод исключения.

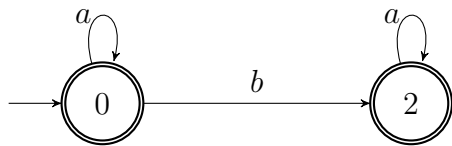


Опишем заключающее состояние в вершине 8: $R = a^*ba^*ab(a + b)^+$.

Уберём вершину 8, опишем состояние в вершине 4: $R = a^*ba^*a (\varepsilon + b(a+b)^+)$.



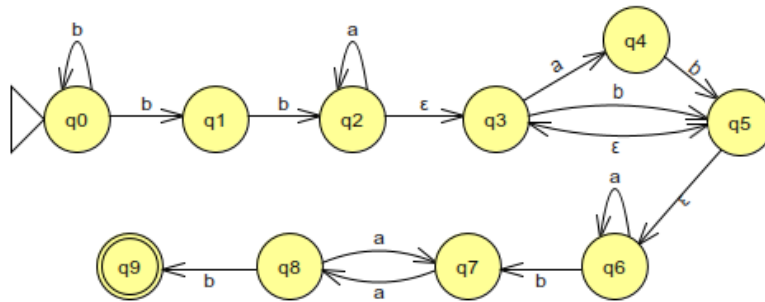
Уберём вершины 4, 6, опишем состояние в вершине 2:
 $R = a^*ba^* (\varepsilon + b(a+b)^* + a(\varepsilon + b(a+b)^+))$.



Данный автомат можно описать выражением $\varepsilon + a^* + a^*b + a^*ba^* = a^*(b + \varepsilon)a^*$.

Ответ: $R = a^*(b + \varepsilon)a^* (\varepsilon + b(a+b)^* + a(\varepsilon + b(a+b)^+))$.

- (vi) Построить ε -НКА, распознающий L_1L_2 .
 $q_0 = S$, остальные состояния пронумерованы от 1 до 9 ($q_1 - q_9$).



- (vii) Детерминизировать один из построенных ε -НКА. Верхние индексы для удобства.

$\delta' \mid$	S^0	$S1^1$	$S12356^2$	23456^3	$S123567^4$	3567^5	234568^6	468^7
$a \mid$	—	—	23456^3	23456^3	234568^6	468^7	234567^8	67^{10}
$b \mid$	$S1^1$	$S12356^2$	$S123567^4$	3567^5	$S123567^4$	3567^5	35679^9	35679^9

$\delta' \mid$	234567^8	35679^9	67^{10}	68^{11}	7^{12}	79^{13}	8^{14}	9^{15}
$a \mid$	234568^6	468^7	68^{11}	67^{10}	8^{14}	8^{14}	—	—
$b \mid$	3567^5	3567^5	7^{12}	79^{13}	—	—	9^{15}	—

Задание 1

- (vi) Доказать, что $L(M^{ND}) = L(M^D) = L$.

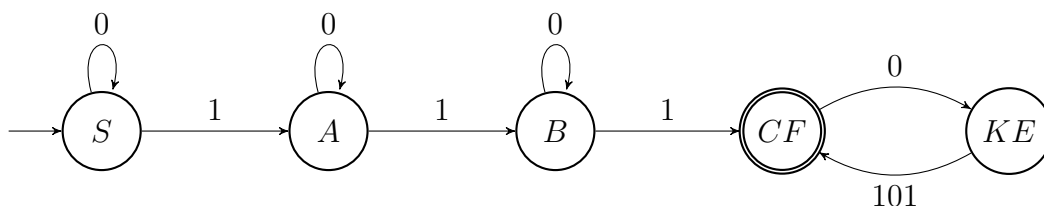
Так как мы получили M^D путём детерминизации M^{ND} , то по теореме о редукции языки этих автоматов эквивалентны. Построение M^{ND} производилось по регулярной грамматике, и по теореме Клини следует, что $L(G) = L(M^{ND})$. Равенство же $L(G) = L$ было доказано во втором пункте первого задания. Таким образом тождество доказано.

Фактическое совпадение языков можно продемонстрировать следующим способом: в пункте (iii) мы вывели регулярное выражение для грамматики, а в пункте (viii) методом исключения получили регулярное выражение, описывающее язык автомата. Так как выражения совпадают, языки эквивалентны.

- (vii) Минимизировать полученный конечный автомат, распознающий язык L , или доказать его минимальность.

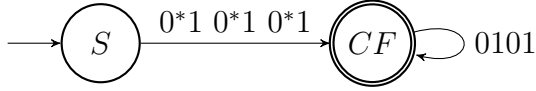
Рассмотрим M^D . Конечные состояния CF и KE являются эквивалентными, поэтому их можно совместить. Все остальные состояния различимы.

Отобразим этот автомат на диаграмме. (101 — два состояния с тремя дугами.)

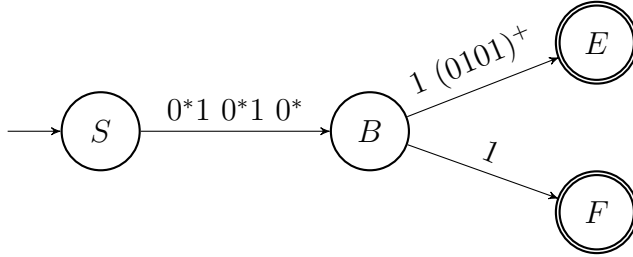


- (viii) Методом последовательного исключения состояний выписать регулярные выражения для $L(M^{ND})$, $L(M^D)$.

M^D : Петли с нулём заменим на 0^* , а цикл $CF-KE$ на $(1010)^*$, $L(M^D) = 0^*10^*10^*1(0101)^*$;

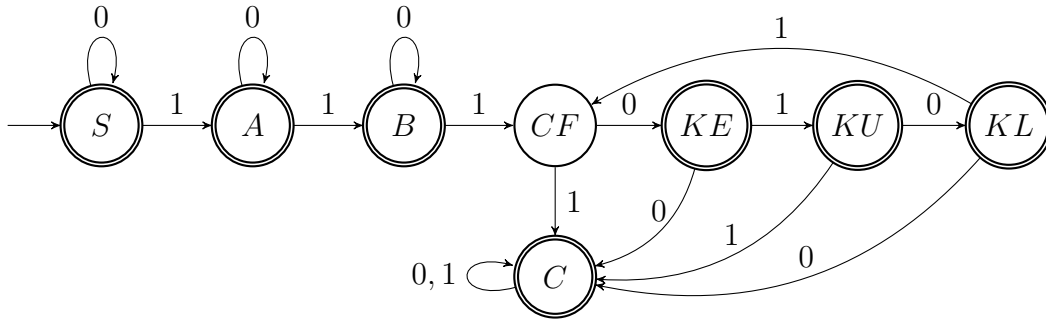


M^{ND} : Стянем вершину C , после чего разобьём SB на два пути: SE, SF .
 $L(M^{ND}) = (0^*10^*10^*1) + (0^*10^*10^*1(0101)^+) = 0^*10^*10^*1(0101)^*$.



- (ix) Построить ДКА, распознающий дополнение $!L$ к языку L , записать $!L$ в виде регулярного выражения.

Доопределим автомат до полного, преобразуем его в дополнение.



Построим грамматику, порождающую этот автомат.

$$\begin{aligned}
 S &\rightarrow \varepsilon \mid 0S \mid 1A, \\
 A &\rightarrow \varepsilon \mid 0A \mid 1B, \\
 B &\rightarrow \varepsilon \mid 0B \mid 1CF, \\
 CF &\rightarrow 0KE \mid 1C, \\
 KE &\rightarrow \varepsilon \mid 1KU \mid 0C, \\
 KU &\rightarrow \varepsilon \mid 0KL \mid 1C, \\
 KL &\rightarrow \varepsilon \mid 1CF \mid 0C, \\
 C &\rightarrow \varepsilon \mid 1C \mid 0C;
 \end{aligned} \tag{3}$$

Решим систему уравнений:

$$\begin{aligned}
 KE &= 101CF + 1 + 10 + (0 + 11 + 100)(0 + 1)^* + \varepsilon, \\
 CF &= 0101CF + 0 + 01 + 010 + (1 + 00 + 011 + 0100)(0 + 1)^* + \varepsilon, \\
 CF &= (0101)^*(0 + 01 + 010 + (1 + 00 + 011 + 0100)(0 + 1)^* + \varepsilon), \\
 S &= 0^*(\varepsilon + 10^*(\varepsilon + 10^*(\varepsilon + 1CF))) = 0^* + 0^*10^* + 0^*10^*1CF
 \end{aligned}$$

Ответ: $0^* + 0^*10^* + 0^*10^*1(0101)^*(0 + 01 + 010 + (1 + 00 + 011 + 0100)(0 + 1)^* + \varepsilon)$.