

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ им. В. Г. ШУХОВА»
(БГТУ им. В.Г. Шухова)



ИНСТИТУТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И УПРАВЛЯЮЩИХ СИСТЕМ

Лабораторная работа №3

по дисциплине: Теория автоматов и формальных языков
тема: «Регулярные языки и конечные распознаватели»

Выполнил: ст. группы ПВ-223
Пахомов Владислав Андреевич

Проверили:
ст. пр. Рязанов Юрий Дмитриевич

Белгород 2024 г.

Лабораторная работа №3
Регулярные языки и конечные распознаватели
Вариант 8

Цель работы: изучить основные способы задания регулярных языков, способы построения, алгоритмы преобразования, анализа и реализации конечных распознавателей.

1. Язык L_1 в алфавите $\{0, 1\}$, представляющий собой множество цепочек, в которых на предпоследнем месте стоит единица, задан грамматикой:

$S \rightarrow A10$

$A \rightarrow A011$

$A \rightarrow 0A$

$A \rightarrow 1A$

$A \rightarrow \varepsilon$

Построить детерминированный конечный распознаватель языка L_1 .

Преобразуем заданную грамматику к автоматной правосторонней. Сейчас она является КС-грамматикой.

Приведём грамматику и устраним левую рекурсию.

Лишних символов в грамматике нет.

В грамматике есть ε -правило. Исключим его.

$S \rightarrow A10$

$S \rightarrow 10$

$A \rightarrow A011$

$A \rightarrow 011$

$A \rightarrow 0A$

$A \rightarrow 0$

$A \rightarrow 1A$

$A \rightarrow 1$

Цепных правил в грамматике нет.

В грамматике есть левая рекурсия. Исключим её.

$S \rightarrow A10$

$S \rightarrow 10$

$A \rightarrow 011B$

$A \rightarrow 0AB$

$A \rightarrow 0B$

$A \rightarrow 1AB$

$A \rightarrow 1B$

$A \rightarrow 011$

$A \rightarrow 0A$

$A \rightarrow 0$

$A \rightarrow 1A$

$A \rightarrow 1$

$B \rightarrow 011B$

$B \rightarrow \varepsilon$

В грамматике есть ε -правило. Исключим его.

$S \rightarrow A10$
 $S \rightarrow 10$
 $A \rightarrow 011$
 $A \rightarrow 011B$
 $A \rightarrow 0A$
 $A \rightarrow 0AB$
 $A \rightarrow 0$
 $A \rightarrow 0B$
 $A \rightarrow 1A$
 $A \rightarrow 1AB$
 $A \rightarrow 1$
 $A \rightarrow 1B$
 $A \rightarrow 011$
 $A \rightarrow 0A$
 $A \rightarrow 0$
 $A \rightarrow 1A$
 $A \rightarrow 1$
 $B \rightarrow 011B$
 $B \rightarrow 011$

Исключим правила-дубликаты:

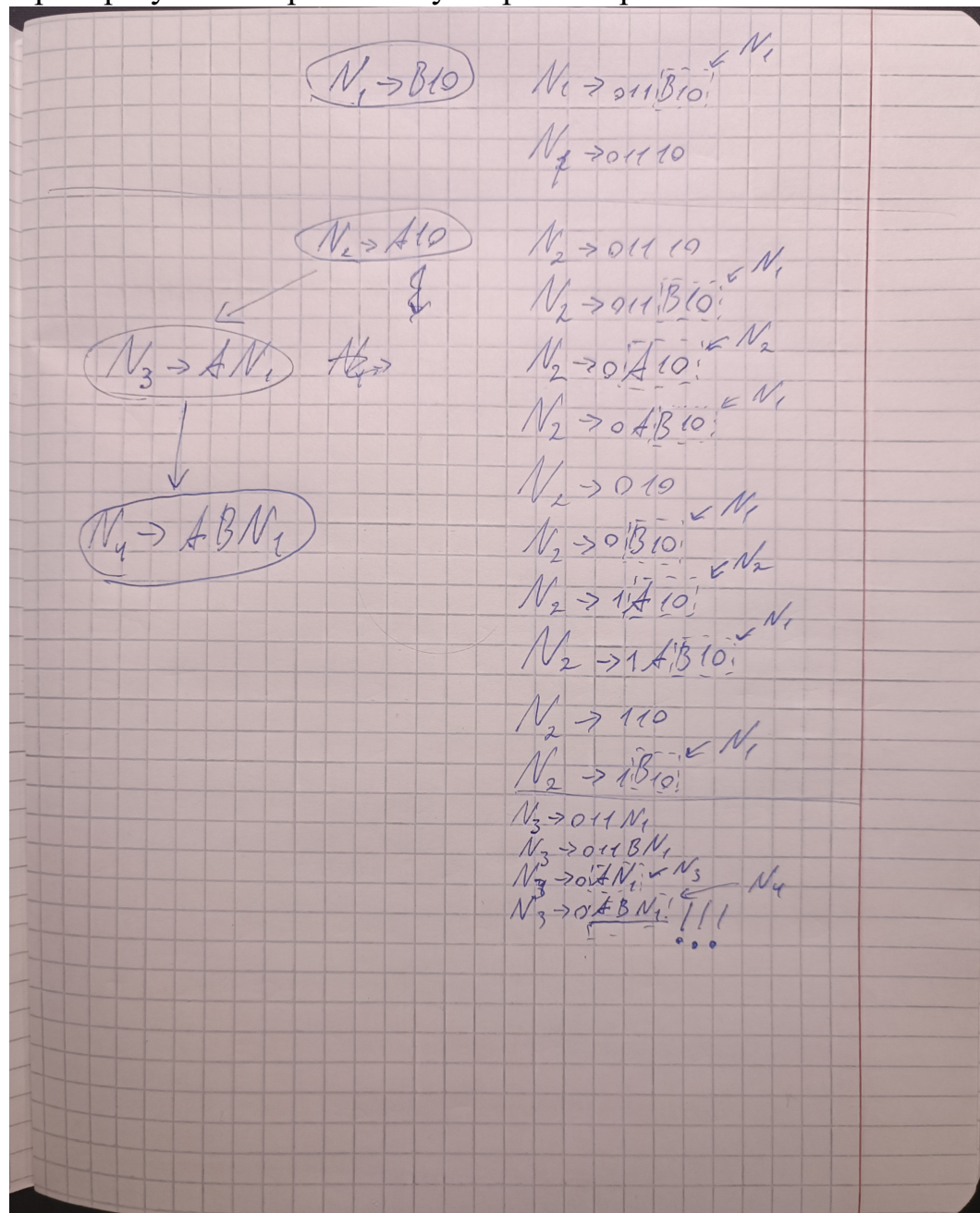
$S \rightarrow A10$
 $S \rightarrow 10$
 $A \rightarrow 011$
 $A \rightarrow 011B$
 $A \rightarrow 0A$
 $A \rightarrow 0AB$
 $A \rightarrow 0$
 $A \rightarrow 0B$
 $A \rightarrow 1A$
 $A \rightarrow 1AB$
 $A \rightarrow 1$
 $A \rightarrow 1B$
 $B \rightarrow 011B$
 $B \rightarrow 011$

Грамматика приведена, а также в ней нет левой рекурсии. Преобразуем грамматику к такому виду, что каждое правило будет начинаться с терминала:

$S \rightarrow 01110$
 $S \rightarrow 011B10$
 $S \rightarrow 0A10$
 $S \rightarrow 0AB10$
 $S \rightarrow 010$
 $S \rightarrow 0B10$
 $S \rightarrow 1A10$
 $S \rightarrow 1AB10$
 $S \rightarrow 110$

$S \rightarrow 1B10$
 $S \rightarrow 10$
 $A \rightarrow 011$
 $A \rightarrow 011B$
 $A \rightarrow 0A$
 $A \rightarrow 0AB$
 $A \rightarrow 0$
 $A \rightarrow 0B$
 $A \rightarrow 1A$
 $A \rightarrow 1AB$
 $A \rightarrow 1$
 $A \rightarrow 1B$
 $B \rightarrow 011B$
 $B \rightarrow 011$

Преобразуем КС-грамматику к правосторонней:



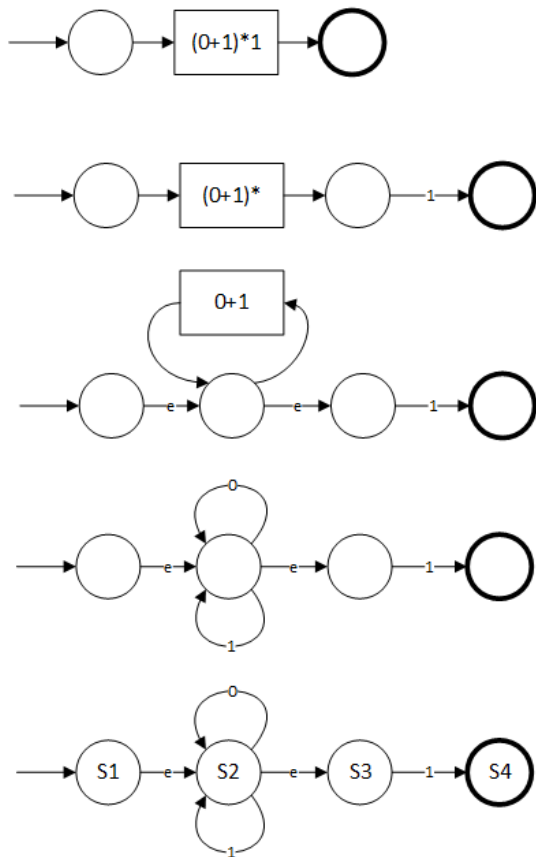
Преобразовать грамматику к правосторонней невозможно, так как в ходе преобразований получили правило (подчёркнутое с !!! в вычислениях) $N_4 \rightarrow ABN_1$. С правилом $N_3 \rightarrow AN_1$ они имеют общий префикс и постфикс, в дальнейшем мы будем получать правила вида AB^*N_1 , получаем рекурсию, и следовательно правостороннюю грамматику с конечным числом правил получить нельзя. Задание невыполнимо.

2. Язык L_2 в алфавите $\{0, 1\}$, представляющий собой множество цепочек, в которых на последнем месте стоит единица, задан регулярным выражением:
 $(0+1)^*1$

Построить детерминированный конечный распознаватель языка L_2 .

Для начала построим конечный недетерминированный распознаватель языка:

Получение
 недетерминированного
 конечного распознавателя:



Данный распознаватель языка не является детерминированным, так как он содержит ϵ -переходы. Преобразуем данный конечный распознаватель языка в детерминированный:

	\downarrow			1
	S1	S2	S3	S4
1		S2	S4	
0		S2		
ϵ	S2	S3		

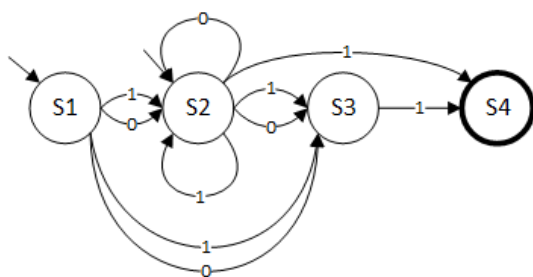
Удалим ϵ -переходы:

ϵ -замыкания: $\epsilon(S1) = \{S1, S2\}$, $\epsilon(S2) = \{S2, S3\}$, $\epsilon(S3) = \{S3\}$, $\epsilon(S4) = \{S4\}$

	↓	↓		1
	$\epsilon(S1)$ {S1, S2}	$\epsilon(S2)$ {S2, S3}	$\epsilon(S3)$ {S3}	$\epsilon(S4)$ {S4}
1	$\epsilon(S2)$	$\epsilon(S2), \epsilon(S4)$	$\epsilon(S4)$	
0	$\epsilon(S2)$	$\epsilon(S2)$		

	↓	↓		1
	S1	S2	S3	S4
1	S2, S3	S2, S3, S4	S4	
0	S2, S3	S2, S3		

Устранение ϵ -переходов



Преобразуем недетерминированный конечный распознаватель в детерминированный:

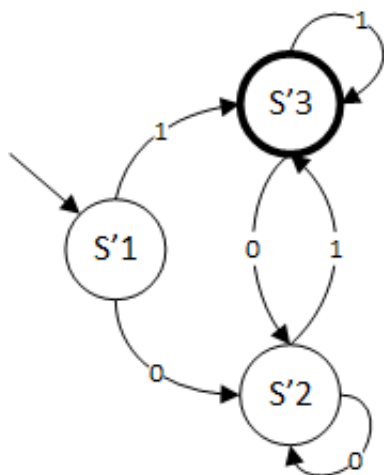
	{S1, S2}	{S2, S3}	{S2, S3, S4}
1	{S2, S3, S4}	{S2, S3, S4}	{S2, S3, S4}
0	{S2, S3}	{S2, S3}	{S2, S3}

Обозначим множества состояний как S'1, S'2, S'3...

S'1 обозначим как начальное состояние, согласно алгоритму, а S'3 обозначим как допускающее состояние, так как множество {S2, S3, S4} включает в себя допускающее состояние S4.

	↓		1
	S'1	S'2	S'3
1	S'3	S'3	S'3
0	S'2	S'2	S'2

Переход к
детерминированному
распознавателю

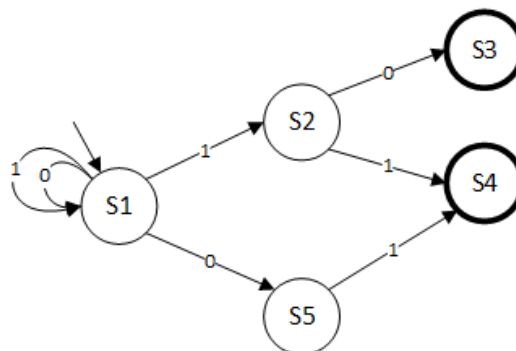


Построили детерминированный конечный распознаватель языка L_2 .

3. Построить минимальный детерминированный конечный распознаватель языка L_3 в алфавите $\{0,1\}$, представляющий собой множество цепочек, в которых хотя бы на одной из последних двух позиций стоит единица.

Пусть у нас будет исходный распознаватель языка L_3 . В начальном состоянии S1 мы итеративно получаем 0 и 1, для окончания работы переходим в состояние S2 по входному символу 1, из него можем попасть в допускающие состояния S3 или S4 по входным символам 0 и 1 соответственно, так как если 1 - предпоследний символ, то строку можем закончить либо 1 либо 0. Если же предпоследний символ - 0, то из состояния S1 можно перейти в состояние S5 при входном символе 0. Однако из S5 мы теперь можем попасть только в S3, так как если предпоследний символ - 0, то последним обязательно должен быть 1. Получили недетерминированный конечный алгоритм без ϵ -переходов:

Исходный
недетерминированный
распознаватель



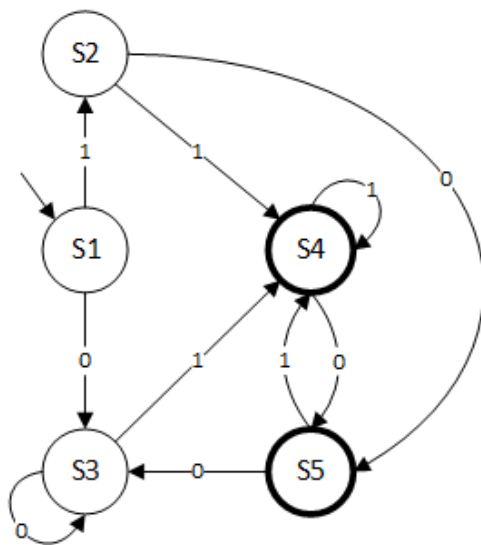
	↓		1	1	
	S1	S2	S3	S4	S5
1	S1, S2	S4			S4
0	S1, S5	S3			

Преобразуем распознаватель в детерминированный:

	{S1}	{S1, S2}	{S1, S5}	{S1, S2, S4}	{S1, S5, S3}
1	{S1, S2}	{S1, S2, S4}	{S1, S2, S4}	{S1, S2, S4}	{S1, S2, S4}
0	{S1, S5}	{S1, S5, S3}	{S1, S5}	{S1, S5, S3}	{S1, S5}

	↓			1	1
	S1	S2	S3	S4	S5
1	S2	S4	S4	S4	S4
0	S3	S5	S3	S5	S3

Детерминированный
распознаватель



Полученный распознаватель является детерминированным, однако является ли он минимальным?

Вывод: в ходе лабораторной работы изучили основные способы задания регулярных языков, способы построения, алгоритмы преобразования, анализа и реализации конечных распознавателей.