

2 Лабораторная работа № 2. Тема: «Построение конечного автомата по регулярной грамматике»

Цель: - закрепить понятия «регулярная грамматика», «недетерминированный и детерминированный конечный автомат»;

- сформировать умения и навыки построения конечного автомата по регулярной грамматике и преобразования недетерминированного конечного автомата к детерминированному конечному автомату.

Основы теории

Распознавателем для регулярной грамматики является конечный автомат (КА).

Определение 2.1. Детерминированным конечным автоматом (ДКА) называется пятерка объектов:

$$M = (Q, T, F, H, Z), \quad (2.1)$$

где Q - конечное множество состояний автомата;

T - конечное множество допустимых входных символов;

F - функция переходов, отображающая множество $Q \times T$ во множество Q ;

H - конечное множество начальных состояний автомата;

Z - множество заключительных состояний автомата, $Z \subseteq Q$.

Определение 2.2. Недетерминированным конечным автоматом (НКА) называется конечный автомат, в котором в качестве функции переходов используется отображение $Q \times T$ во множество всех подмножеств множества состояний автомата $P(Q)$, т.е. функция переходов неоднозначна, так как текущей паре (q, t) соответствует множество очередных состояний автомата $q' \in P(Q)$.

Способы представления функции переходов

Командный способ. Каждую команду КА записывают в форме $F(q, t) = p$, где $q, p \in Q, t \in T$.

Табличный способ. Строки таблицы переходов соответствуют входным символам автомата $t \in T$, а столбцы – состояниям Q . Ячейки таблицы заполняются новыми состояниями, соответствующими значению функции $F(q, t)$. Неопределенным значениям функции переходов соответствуют пустые ячейки таблицы.

Графический способ. Строится диаграмма состояний автомата – неупорядоченный ориентированный помеченный граф. Вершины графа помечены именами состояний автомата. Дуга ведет из состояния q в состояние p и помечается списком всех символов $t \in T$, для которых $F(q, t) = p$. Вершина, соответствующая входному состоянию автомата, снабжается стрелкой. Заклю-

тельное состояние на графе обозначается двумя концентрическими окружностями.

Алгоритм 2.1. Построение КА по регулярной грамматике

Вход: регулярная грамматика $G = (V_T, V_N, P, S)$.

Выход: КА $M = (Q, T, F, H, Z)$.

Шаг 1. Пополнить грамматику правилом $A \rightarrow aN$, где $A \in V_N, a \in V_T$ и N - новый нетерминал, для каждого правила вида $A \rightarrow a$, если в грамматике нет соответствующего ему правила $A \rightarrow aB$, где $B \in V_N$.

Шаг 2. Начальный символ грамматики S принять за начальное состояние КА H . Из нетерминалов образовать множество состояний автомата $Q = V_N \cup \{N\}$, а из терминалов – множество символов входного алфавита $T = V_T$.

Шаг 3. Каждое правило $A \rightarrow aB$ преобразовать в функцию переходов $F(A, a) = B$, где $A, B \in V_N, a \in V_T$.

Шаг 4. Во множество заключительных состояний включить все вершины, помеченные символами $B \in V_N$ из правил вида $A \rightarrow aB$, для которых имеются соответствующие правила $A \rightarrow a$, где $A, B \in V_N, a \in V_T$.

Шаг 5. Если в грамматике имеется правило $S \rightarrow \varepsilon$, где S - начальный символ грамматики, то поместить S во множество заключительных состояний.

Шаг 6. Если получен НКА, то преобразовать его в ДКА.

Алгоритм 2.2. Преобразование НКА в ДКА

Вход: НКА $M = (Q, T, F, H, Z)$.

Выход: ДКА $M' = (Q', T, F', H, Z')$.

Шаг 1. Пометить первый столбец таблицы переходов M' ДКА начальным состоянием (множеством начальных состояний) НКА M .

Шаг 2. Заполняем очередной столбец таблицы переходов M' , помеченный символами D , для этого определяем те состояния M , которые могут быть достигнуты из каждого символа строки D при каждом входном символе x . Поместить каждое найденное множество R (в том числе \emptyset) в соответствующие позиции столбца D таблицы M' , т.е.:

$$F'(D, x) = \{s \mid s \in F(t, x) \text{ для некоторого } t \in D\}.$$

Шаг 3. Для каждого нового множества R (кроме \emptyset), полученного в столбце D таблицы переходов M' , добавить новый столбец в таблицу, помеченный R .

Шаг 4. Если в таблице переходов КА M' есть столбец с незаполненными позициями, то перейти к шагу 2.

Шаг 5. Во множество Z' ДКА M' включить каждое множество, помещающее столбец таблицы переходов M' и содержащее $q \in Z$ НКА M .

Шаг 6. Составить таблицу новых обозначений множеств состояний и определить ДКА M' в этих обозначениях.

Пример 2.1. Дана регулярная грамматика $G = (\{a, b\}, \{S, A, B\}, P, S)$ с правилами P : 1) $S \rightarrow aB \mid aA$; 2) $B \rightarrow bB \mid a$; 3) $A \rightarrow aA \mid b$. Построить по регулярной грамматике КА и преобразовать полученный автомат к детерминированному виду.

Решение задачи включает следующую последовательность действий.

1 Построим по регулярной грамматике КА.

1.1 Пополним грамматику правилами $A \rightarrow bN$ и $B \rightarrow aN$, где N - новый нетерминал.

1.2 Начальное состояние конечного автомата $H = S$. Множество состояний автомата $Q = V_N = \{S, A, B, N\}$, множество символов входного алфавита $T = V_T = \{a, b\}$.

1.3 Значения сформированной функции переходов даны в таблице 2.1.

Таблица 2.1 – Функция переходов автомата M

F	S	A	B	N
a	A, B	A	N	\emptyset
b	\emptyset	N	B	\emptyset

1.4 Множество заключительных состояний $Z = \{N\}$.

1.5 Для начального символа грамматики ϵ -правила отсутствуют.

Конечный автомат M - недетерминированный, граф НКА представлен на рисунке 2.1 слева.

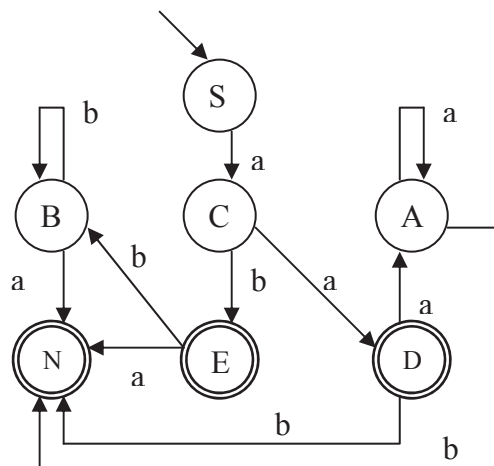
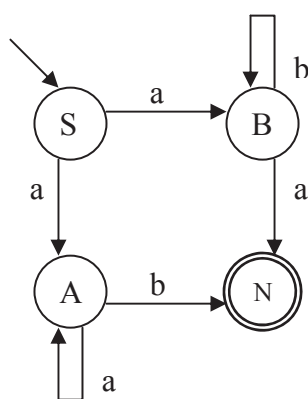


Рисунок 2.1 - Граф НКА (слева) и ДКА (справа) для P-грамматики

2 Построим по НКА M ДКА M' .

2.1 Строим таблицу переходов для ДКА M' (таблица 2.2).

Таблица 2.2 – Построение функции переходов для ДКА M'

Шаг	1	2	3	4	5	6	7
F	S	A, B	A, N	B, N	A	N	B
a	A, B	A, N	A	N	A	\emptyset	N
b	\emptyset	B, N	N	B	N	\emptyset	B

2.2 Во множество заключительных состояний автомата M' включим элементы $Z' = \{(A, N), (B, N), N\}$.

2.3 Введем следующие новые обозначения состояний автомата M' : $(A, B)=C$, $(A, N)=D$, $(B, N)=E$.

2.4 Искомый ДКА определяется следующей пятеркой объектов: $Q' = \{S, A, B, C, D, E, N\}$, $T = \{a, b\}$, функция переходов задана таблицей 2.3, $H = \{S\}$, $Z' = \{N, D, E\}$.

Граф полученного ДКА представлен на рисунке 2.1 справа.

Таблица 2.3 – Функция переходов для ДКА M'

F'	S	A	B	C	D	E	N
a	C	A	N	D	A	N	\emptyset
b	\emptyset	N	B	E	N	B	\emptyset

Постановка задачи к лабораторной работе № 2

Разработать программное средство, реализующее следующие функции:

- 1) ввод произвольной формальной грамматики с клавиатуры и проверка ее на принадлежность к классу регулярных грамматик;
- 2) построение по заданной регулярной грамматике конечного автомата;
- 3) преобразование недетерминированного конечного автомата к детерминированному конечному автомату;
- 4) вывод графа результирующего конечного автомата на экран.

Варианты индивидуального задания представлены в таблице 2.4.

Таблица 2.4 – Варианты индивидуального задания к лабораторной работе № 2

Вариант	Регулярная грамматика
1	$G = (\{S, C, D\}, \{0, 1\}, P, S)$, где P : 1) $S \rightarrow 1C \mid 0D$; 2) $C \rightarrow 0D \mid 0S \mid 1$; 3) $D \rightarrow 1C \mid 1S \mid 0$.
2	$G = (\{S, A, B, C\}, \{a, b, c\}, P, S)$, где P : 1) $S \rightarrow aA \mid bB \mid aC$; 2) $A \rightarrow bA \mid bB \mid c$; 3) $B \rightarrow aA \mid cC \mid b$; 4) $C \rightarrow bB \mid bC \mid a$.

Продолжение таблицы 2.4 – Варианты индивидуального задания к лабораторной работе № 2

Вариант	Регулярная грамматика
3	$G=(\{K, L, M, N\}, \{a, b, +, -, \perp\}, P, K)$, где P : 1) $K \rightarrow aL \mid bM$; 2) $L \rightarrow -N \mid -M$; 3) $M \rightarrow +N$; 4) $N \rightarrow aL \mid bM \mid \perp$.
4	$G=(\{X, Y, Z, W, V\}, \{0, 1, \sim, \#, \&\}, P, X)$, где P : 1) $X \rightarrow 0Y \mid 1Z \mid \varepsilon$; 2) $Y \rightarrow 0Z \mid \sim W \mid \#$; 3) $Z \rightarrow 1Y \mid 1W \mid 0V$; 4) $W \rightarrow 0W \mid 1W \mid \#$; 5) $V \rightarrow \&Z$.
5	$G=(\{K, L, M, N, Q, P, R, S\}, \{0, 1, *, \$, /\}, V, K)$, где V : 1) $K \rightarrow 1L \mid 0N$; 2) $L \rightarrow 0M \mid 0P \mid /Q$; 3) $N \rightarrow 1R \mid 1M \mid *S$; 4) $Q \rightarrow 1P$; 5) $P \rightarrow *L \mid \$$; 6) $M \rightarrow \$$; 7) $S \rightarrow 0R$; 8) $R \rightarrow /N \mid \$$.
6	$G=(\{E, A, B, C, D\}, \{0, 1, a, b, c\}, P, E)$, где P : 1) $E \rightarrow 0A \mid \varepsilon$; 2) $A \rightarrow aB \mid aD$; 3) $B \rightarrow bB \mid 1C \mid c$; 4) $D \rightarrow aD \mid 0C \mid c$.
7	$G=(\{X, Y, Z, V, W\}, \{0, 1, x, y, z\}, P, X)$, где P : 1) $X \rightarrow yY \mid zZ$; 2) $Y \rightarrow 1V$; 3) $Z \rightarrow 0W \mid 0Y$; 4) $V \rightarrow xZ \mid xW \mid 1$; 5) $W \rightarrow 1Y \mid 0$.
8	$G=(\{S, A, B, C, D\}, \{a, b, c, d, \perp\}, P, S)$, где P : 1) $S \rightarrow aA \mid bB$; 2) $A \rightarrow cC \mid \perp$; 3) $C \rightarrow cC \mid cA$; 4) $B \rightarrow dD \mid \perp$; 5) $D \rightarrow dD \mid dB$.
9	$G=(\{K, L, M, N, P\}, \{0, 1, \&, \%, a, b\}, C, K)$, где C : 1) $K \rightarrow 1M \mid \varepsilon$; 2) $M \rightarrow 0L \mid \&N \mid \&P$; 3) $L \rightarrow 1L \mid 0L \mid \%P$; 4) $N \rightarrow aN \mid bN \mid \%P$; 5) $P \rightarrow 1P \mid aP \mid 0$.
10	$G=(\{I, J, K, M, N\}, \{0, 1, \sim, !\}, P, I)$, где P : 1) $I \rightarrow 0J \mid 1K \mid 0M$; 2) $J \rightarrow \sim K \mid 0M$; 3) $K \rightarrow \sim M \mid 0J \mid 0N$; 4) $M \rightarrow 1K \mid !$; 5) $N \rightarrow 0I \mid 1I \mid !$.
11	$G=(\{S, A, B, C, D, E\}, \{a, b, c, d, e, \$, \perp\}, P, S)$, где P : 1) $S \rightarrow aA \mid bB \mid cC$; 2) $A \rightarrow dD$; 3) $B \rightarrow \#D \mid \$E$; 4) $D \rightarrow dD \mid dB \mid \perp$; 5) $C \rightarrow cE$; 6) $E \rightarrow eE \mid eB \mid \perp$.
12	$G=(\{X, Y, Z, V\}, \{(\, , \,), y, z, v\}, P, X)$, где P : 1) $X \rightarrow (Y \mid \varepsilon$; 2) $Y \rightarrow yY \mid zY \mid zZ$; 3) $Z \rightarrow zZ \mid vZ \mid vV$; 4) $V \rightarrow vV \mid)$.