2 КОНЕЧНЫЕ АВТОМАТЫ

Конечным автоматом называется набор

$$< \Sigma,\ Q,\ q_0,\ Q_{\text{закл}},\ \Phi >,$$

где Σ является конечным множеством и называется *входным алфавитом*, Q является конечным множеством и называется *алфавитом состояний*, $q_0 \in Q$ называется *начальным состоянием*, элементы из $Q_{\text{закл}} \subseteq Q$ называются заключительными состояниями, а Φ является отображением $(Q \times \Sigma)$ в Q и называется *множеством правил* этого конечного автомата. *Правилами* этого конечного автомата называются тройки $< q, a, q' >$, в которых $q, q' \in Q$, а $a \in \Sigma$. Требуется, чтобы для каждой пары $< q, a >$ из $(Q \times \Sigma)$ имелось одно и только одно такое состояние q', что тройка $< q, a, q' >$ является правилом рассматриваемого конечного автомата. Это условие называется *детерминированностью* этого конечного автомата.

Конечный автомат, например, можно задать следующим образом.

Пример 2.0.1.

	$q1$ *начальное*	$q2$	$q3$	$q4$ *заключительное*	$q5$	$q6$	$q7$	$q8$
0	$q8$	$q3$	$q8$	$q6$	$q6$	$q7$	$q7$	$q8$
1	$q2$	$q8$	$q4$	$q5$	$q5$	$q4$	$q5$	$q8$

Здесь строчки помечены буквами входного алфавита, столбцы помечены состояниями, указаны начальное и заключительные состояния. Для каждой входной буквы и состояния указано на пересечении соответствующих строчки и столбца следующее состояние. Таким способом можно задать правила автомата.

Работа конечного автомата описывается следующим образом. Автомат начинает работу в начальном состоянии и после восприятия очередной буквы переходит согласно с соответствующим правилом в следующее состояние. Если после прохождения слова автомат переходит в одно из заключительных состояний, то слово *воспринимается* этим автоматом.

Например, для слова "101000101"получаем следующую последовательность состояний:

	$q1$ начальное	$q2$	$q3$	$q4$ заключительное	$q5$	$q6$	$q7$	$q8$
0	$q8$	$q3$	$q8$	$q6$	$q6$	$q7$	$q7$	$q8$
1	$q2$	$q8$	$q4$	$q5$	$q5$	$q4$	$q5$	$q8$

	1	0	1	0	0	0	1	0	1
$q1$	$q2$	$q3$	$q4$	$q6$	$q7$	$q7$	$q5$	$q6$	$q4$

Так как последнее состояние оказалось заключительным, это слово воспринимается рассматриваемым автоматом.

Более формально, пусть на вход конечного автомата $M = \langle \Sigma, Q, q_0, Q_{\text{закл}}, \Phi \rangle$ подаётся слово w из Σ^*. Автомат в начальный момент времени считывает первую букву слова w. Сам автомат находится в состоянии q_0.

Такт работы автомата M состоит в том, что он переходит в новое состояние и затем считывает следующую букву входного слова. Если следующей буквы нет, автомат останавливается, в противном случае он делает следующий такт и так далее.

Мы будем иногда использовать и следующую терминологию. Если на вход автомата подаётся слово w, будем говорить, что на входной ленте записано это слово. Будем также говорить, что считываемая буква обозревается на рассматриваемом такте работы автомата.

Пусть автомат M обозревает в данный момент букву a из Σ, а сам автомат находится в состоянии q из Q. Тогда новое состояние q' на этом такте выбирается с условием, что тройка $\langle q, a, q' \rangle$ принадлежит множеству Φ правил автомата M.

Легко видеть, что на слове w автомат M совершит ровно $|w|$ тактов. Состояние, в котором был автомат на последнем из этих тактов и на котором он остановился, будем называть последним состоянием.

Будем говорить, что автомат M воспринимает слово w в алфавите Σ, если последним состоянием M на слове w будет состояние, принадлежащее множеству $Q_{\text{закл}}$.

Пусть автомат M находится в состоянии q, обозревает букву a на входной ленте, и справа от обозреваемой буквы на ленте записано слово w.

Определение 2.0.1 (работы конечного автомата). *Если $q' \in Q$ такое, что $\langle q, a, q' \rangle \in \Phi$, то будем писать*

$$(q, aw) \mid\Rightarrow_M (q', w)$$

и

$$(q, aw) \mid\Rightarrow^1_M (q', w).$$

Пусть

$$(q, aw'w) \mid\Rightarrow^{n+1}_M (q', w)$$

означает, что существует такое состояние q'' автомата M, что

$$(q, aw'w) \mid\Rightarrow_M (q'', w'w)$$

и

$$(q'', w'w) \mid\Rightarrow^n_M (q', w).$$

Отношение

$$(q, w'w) \mid\Rightarrow^+_M (q', w)$$

означает, что

$$(q, w'w) \mid\Rightarrow^n_M (q', w)$$

для некоторого целого положительного n.

Автомат M воспринимает слово $w \in \Sigma^*$ тогда и только тогда, когда существует такое $q \in Q_{\text{закл}}$, что

$$(q_0, w) \mid\Rightarrow^+_M (q, \lambda),$$

где λ — пустое слово. Автомат M воспринимает пустое слово тогда и только тогда, когда $q_0 \in Q_{закл}$.

Определение 2.0.2 (языка конечного автомата). *Языком конечного автомата $M = \langle \Sigma,\ Q,\ q_0,\ Q_{закл},\ \Phi \rangle$ называется множество всех слов в алфавите Σ, которые воспринимаются автоматом M. Язык автомата M мы будем обозначать через $L(M)$. Мы будем также говорить, что язык $L(M)$ распознается (или задается) автоматом M.*

Покажем, что в примере 2.0.1 автомат задаёт язык, состоящий из тех и только тех слов, составленных из букв 0 и 1, которые начинаются и оканчиваются на "101".

	$q1$ начальное	$q2$	$q3$	$q4$ заключительное	$q5$	$q6$	$q7$	$q8$
0	$q8$	$q3$	$q8$	$q6$	$q6$	$q7$	$q7$	$q8$
1	$q2$	$q8$	$q4$	$q5$	$q5$	$q4$	$q5$	$q8$

В самом деле, этот автомат сначала проверяет, начинается ли слово на 101. Если не начинается, автомат переходит в незаключительное состояние $q8$, в котором и остаётся уже навсегда. Иначе автомат переходит в заключительное состояние $q4$, так как слово 101 заканчивается на 101 и должно восприниматься автоматом.

После этого состояние $q4$ говорит, что слово оканчивается на 101, состояние $q5$ говорит, что слово оканчивается на 11 или на 001, состояние $q6$ говорит, что слово оканчивается на 10, а состояние $q7$ говорит, что слово оканчивается на 00.

Действительно, если это верно для начинающихся на 101 слов длины k, то это верно и для таких слов длины $k+1$.

Если мы находимся в состоянии $q4$ и видим 0, то прочитанное слово заканчивается на 10 и мы переходим, как и должно быть, в $q6$. Если же мы видим 1, то переходим в $q5$, так как прочитанное слово теперь оканчивается на 11. Точно так же ведёт себя состояние $q5$. В состоянии $q6$, обозревая 0, мы переходим в $q7$, так как слово теперь оканчивается на 00, а, обозревая 1, переходим в $q4$, так как теперь слово оканчивается на 101. В состоянии $q7$, обозревая 0, переходим в $q7$, так как прочитанное слово снова оканчивается на 00, а, обозревая 1, переходим в $q5$, так как прочитанное слово оканчивается на 001.

Итак, мы переходим в заключительное состояние тогда и только тогда, когда прочитанное слово оканчивается на 101. Поэтому автомат задаёт язык, состоящий из слов, начинающихся на 101 и оканчивающихся на 101.