

8. Автоматы и преобразователи

Процесс распознавания и переработки цепочки символов можно рассматривать как процесс функционирования некоторого условного прибора.

Распознающий автомат имеет:

входную ленту,
устройство чтения,
устройство управления с конечной памятью,
вспомогательную память.

Текущее состояние распознавателя называют **конфигурацией**.

Для каждой конфигурации определены:

состояние устройства управления (их конечное множество),
положение устройства чтения,
состояние вспомогательной памяти.

В множестве конфигураций выделяются начальная конфигурация и конечная конфигурация. Переход от одной конфигурации к другой осуществляется по тактам.

Распознаватель допускает входную цепочку, если он, обрабатывая эту цепочку начиная с начальной конфигурации, переходит в конечную конфигурацию за конечное число тактов.

Конечный автомат — простейший распознаватель без вспомогательной памяти.

Используют следующее определение конечного автомата: $K = (Q, T, \delta, q_0, F)$, где

Q — конечное множество состояний устройства управления,

T — алфавит входных символов,

δ — функция переходов (отображение $Q \times T \rightarrow Q$),

$q_0 \in Q$ — начальное состояние,

$F \subset Q$ — множество заключительных состояний.

Если функция δ — однозначная, то КА называют детерминированным. Если функция δ — многозначная, то КА называют недетерминированным.

Конфигурация автомата $(q, w) \in Q \times T^*$, при этом начальная конфигурация (q_0, w) и конечная конфигурация $(q, \varepsilon) \mid q \in F$. Здесь w — цепочка символов, которые еще не были обработаны.

Переходы от конфигурации к конфигурации обозначаются знаком \vdash .

Распознаватель $K = (Q, T, \delta, q_0, F)$ допускает входную цепочку $w \in T^*$, если $(q_0, w) \vdash (q, \varepsilon)$, $q \in F$.

Язык, определяемый конечным автоматом K ,

$$L(K) = \{w \in T^* \mid (q_0, w) \vdash (q, \varepsilon), q \in F\}$$

Пример. Конечный автомат, допускающий цепочки из 0 и 1, в которых имеется подцепочка 11.

$$K = (\{q_0, q_1, q_2\}, \{0, 1\}, \delta, q_0, \{q_2\})$$

Функция переходов:

$$\begin{aligned} \delta(q_0, 0) &= \{q_0\}, & \delta(q_0, 1) &= \{q_1\}, & \delta(q_1, 0) &= \{q_0\}, \\ \delta(q_1, 1) &= \{q_2\}, & \delta(q_2, 0) &= \{q_2\}, & \delta(q_2, 1) &= \{q_2\} \end{aligned}$$

Часто функция переходов задается как таблица переходов или как диаграмма переходов.

Доказано, что **множество языков, допускаемых конечными автоматами, совпадает с множеством языков, порождаемых автоматными грамматиками.**

Конечный преобразователь анализирует цепочку символов на входной ленте и записывает другую цепочку символов на выходной ленте. По определению $M = (Q, T, D, \delta, q_0, F)$, где

- Q — конечное множество состояний устройства управления,
- T — алфавит входных символов,
- D — алфавит выходных символов,
- δ — функция переходов (отображение $Q \times T \rightarrow Q$),
- $q_0 \in Q$ — начальное состояние,
- $F \subset Q$ — множество заключительных состояний.

Конфигурация конечного преобразователя $(q, x, y) \in Q \times T^* \times D^*$.

Цепочка символов $y \in D^*$ называется выходом для цепочки символов $x \in T^*$, если $(q_0, x, \varepsilon) \vdash (q, \varepsilon, y)$ для некоторого $q \in F$.

Автомат с магазинной памятью (МП-автомат) представляет собой конечный автомат, дополненный неограниченной памятью с доступом только к крайнему символу.

$P = (Q, T, \Gamma, \delta, q_0, Z_0, F)$, где

- Q — конечное множество состояний устройства управления,
- T — алфавит входных символов,
- Γ — алфавит символов магазина,
- δ — функция переходов (отображение $Q \times T \times \Gamma \rightarrow Q \times \Gamma^*$),
- $q_0 \in Q$ — начальное состояние устройства управления,
- $Z_0 \in \Gamma$ — начальный символ в магазине,
- $F \subset Q$ — множество заключительных состояний.

Конфигурация МП-автомата: q — состояние устройства управления, x — необработанная часть входной цепочки и α — содержимое магазина.

Начальная конфигурация (q_0, w, Z_0) , заключительная конфигурация (q, ε, α) .

Такт работы $(q, x, \alpha) \vdash (q', x', \alpha')$.

МП-автомат допускает цепочку символов $w \in T^*$, если $(q_0, w, Z_0) \vdash^* (q, \varepsilon, \alpha)$ при некоторых $q \in F$ и $\alpha \in \Gamma^*$.

Язык, определяемый (допускаемый) МП-автоматом, образуют все распознаваемые им цепочки.

Пример. МП-автомат, допускающий язык $L = \{a^n b^n \mid n \geq 0\}$.

Во время работы МП-автомата операции над крайним символом в магазине не зависят от других символов в магазине.

Расширенные МП-автоматы допускают замену конечной цепочки крайних символов в магазине на другую конечную цепочку.

Множество языков, допускаемых автоматами с магазинной памятью, совпадает с множеством языков, порождаемых контекстно свободными-грамматиками.

Преобразователь с магазинной памятью (МП-преобразователь) представляет собой МП-автомат, имеющий устройство записи символов на выходную ленту.

$D = (Q, T, D, \Gamma, \delta, q_0, Z_0, F)$, где

Q — конечное множество состояний устройства управления,

T — алфавит входных символов,

D — алфавит выходных символов,

Γ — алфавит символов магазина,

δ — функция переходов (отображение $Q \times T \times \Gamma \rightarrow Q \times \Gamma^*$),

$q_0 \in Q$ — начальное состояние устройства управления,

$Z_0 \in \Gamma$ — начальный символ в магазине,

$F \subset Q$ — множество заключительных состояний.

Конфигурация МП-преобразователя: q — состояние устройства управления, x — необработанная часть входной цепочки, α — содержимое магазина и y — цепочка символов на выходной ленте.

Пример. МП-преобразователь, переводящий арифметическое выражение в infixной форме в эквивалентную префиксную форму.