Глава 5

# МАГАЗИННЫЕ АВТОМАТЫ

## § 5.1. Неформальное описание

В этой главе мы рассмотрим простое устройство — магазинный автомат[7] (pda — pushdown automaton), которое адекватно классу КС-языков в том смысле, что любой КС-язык распознается каким-нибудь магазинным автоматом, и любой магазинный автомат распознает некоторый КС-язык.

Магазинный автомат подобен конечному автомату, но в отличие от последнего имеет рабочую память — *магазин*, в который записываются символы из еще одного алфавита — *алфавита магазинных символов*. Каждое движение магазинного автомата определяется в зависимости от текущего *состояния управления*, *входного символа* или независимо от него — так называемые ε-*движения* и от *верхнего символа магазина*. Одно движение магазинного автомата состоит в замещении верхнего символа магазина некоторой магазинной цепочкой, в частности пустой (стирание верхнего символа магазина), и переходе в новое состояние управления. При этом текущим входным символом становится следующий символ на входной ленте, если выполняется движение, зависящее от входного символа, либо текущий входной символ остается тем же самым, если выполняется ε-движение. Поскольку движение зависит от верхнего символа магазина, то с самого начала в магазине находится один символ — *начальный символ магазина*.

Считается, что некоторая цепочка *принята*, если магазинный автомат из начального состояния управления, имея в магазине единственный — начальный символ магазина и прочитав данную цепочку на входе, переходит в одно из своих *конечных состояний* или *опустошает магазин*. Каждый конкретный магазинный автомат использует только какой-нибудь один из этих двух признаков приема входной цепочки. Как и в случае конечных автоматов существуют две модели магазинных автоматов — *недетерминированная* и *детерминированная*. В недетерминированной модели автомат каждый раз имеет возможность некоторого конечного выбора движений и совершает одно из них. Входная цепочка считается принятой, если существует хотя бы одна последовательность выборов движений, которая приводит автомат к *конечной конфигурации* — входная цепочка прочитана, текущее состояние — конечное или (в другом варианте) магазин пуст. В противном случае она не принимается. Множество всех цепочек, принимаемых данным магазинным автоматом, называется *языком*, *распознаваемым этим магазинным автоматом*.

---

[7] Вместо этого термина часто используется сокращение МП-автомат.

В данной главе будет показано, что оба определения приема эквивалентны в том смысле, что если язык принимается некоторым магазинным автоматом при пустом магазине, то он может быть принят некоторым другим магазинным автоматом при конечном состоянии, и наоборот.

Кроме того, будет доказана основная теорема о том, что язык принимается недетерминированным МП-автоматом тогда и только тогда, когда он является КС-языком. Известно, что класс языков, принимаемых детерминированными МП-автоматами, является строгим подклассом языков, принимаемых недетерминированными МП-автоматами.

## § 5.2. Формальное определение

***Определение* 5.1.** Недетерминированный магазинный автомат есть *формальная система* $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, где $Q$ — конечное *множество состояний*; $\Sigma$ — конечный *входной алфавит*; $\Gamma$ — конечный *магазинный алфавит;* $\delta$ — *отображение типа* $Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \to 2^{Q \times \Gamma^*}$, представляющее конечное *управление автомата;* $q_0 \in Q$ *— начальное состояние*; $Z_0 \in \Gamma$ *— начальный символ магазина*, который в самом начале является единственным содержимым магазина; $F \subseteq Q$ *— множество конечных состояний.*

Мы будем придерживаться следующей системы обозначений: строчные буквы из начала латинского алфавита — отдельные входные символы; строчные буквы из конца латинского алфавита служат для обозначения цепочек входных символов; строчные греческие буквы — цепочки магазинных символов; прописные латинские буквы — отдельные магазинные символы.

***Определение* 5.2.** Для описания движений МП-автомата будем использовать понятие *конфигурации*, под которой будем подразумевать тройку $(q, x, \alpha)$, где $q \in Q$ — *текущее состояние управления*; $x \in \Sigma^*$ — *непросмотренная часть входной цепочки* (от текущего символа до ее конца) ; $\alpha \in \Gamma^*$ — *магазинная цепочка*, причем крайний левый ее символ считается находящимся на вершине магазина.

*Начальная конфигурация* есть $(q_0, x, Z_0)$, где $x$ — вся входная цепочка. *Конечная конфигурация* определяется по-разному: в зависимости от того, какой признак приема используется. Если прием определяется при конечном состоянии, то конечная конфигурация есть $(q, \varepsilon, \alpha)$ , где $q \in F$ — автомат достиг конечного состояния; $\varepsilon$ означает, что вся входная цепочка прочитана; $\alpha \in \Gamma^*$ — произвольная магазинная цепочка. Достижение конечного состояния не означает завершения работы автомата, а сигнализирует лишь о том, что прочитанная к этому моменту часть входной цепочки принимается. За время сканирования входной цепочки конечные состояния могут достигаться несколько раз. Если прием определяется при пустом магазине, то конечная конфигурация есть $(q, \varepsilon, \varepsilon)$, где $q \in Q$ — любое состояние; $\varepsilon$ во второй позиции означает, как и в предыдущем случае, что вся входная цепочка прочитана; $\varepsilon$ в третьей позиции означает, что магазин пуст. При пустом магазине никакое дальнейшее движение не определено.

Запись $\delta(q, a, Z) = \{(p_1, \gamma_1), (p_2, \gamma_2), \ldots, (p_m, \gamma_m)\}$, где $q, p_i \in Q$, $a \in \Sigma$, $Z \in \Gamma$, $\gamma_i \in \Gamma^*$ ($i = 1, 2, \ldots, m$), означает, что магазинный автомат в состоянии $q$ с символом $a$ на входе и символом $Z$ на вершине магазина может для любого $i$ перейти в состояние $p_i$, заменить $Z$ на $\gamma_i$ и продвинуться на входе к следующей позиции (при этом крайний левый символ $\gamma_i$ окажется на вершине магазина).

Пусть текущая конфигурация имеет вид: $(q, ax, Z\alpha)$, где $x \in \Sigma^*$. В терминах конфигураций одно из возможных движений представляется следующим образом: $(q, ax, Z\alpha) \vdash (p_i, x, \gamma_i\alpha)$ для любого $i$, $1 \le i \le m$.

Запись $\delta(q, \varepsilon, Z) = \{(p_1, \gamma_1), (p_2, \gamma_2), \ldots, (p_m, \gamma_m)\}$ означает, что pda в состоянии $q$ независимо от того, какой символ на входе, и с символом $Z$ на вершине магазина может для любого $i$ перейти в состояние $p_i$, заменить $Z$ на $\gamma_i$, не изменяя текущей позиции на входе.

Пусть текущая конфигурация имеет вид $(q, x, Z\alpha)$. Тогда в терминах конфигураций одно из возможных движений представляется следующим образом: $(q, x, Z\alpha) \vdash (p_i, x, \gamma_i\alpha)$ для любого $i$, $1 \le i \le m$.

Если $\gamma_i = \varepsilon$, то происходит стирание верхнего символа магазина — вершина магазина опускается; если $|\gamma_i| = 1$, то происходит замена верхнего символа магазина; если $|\gamma_i| > 1$, то вершина магазина поднимается.

Таким образом, мы ввели на множестве конфигураций pda отношение непосредственного следования одной конфигурации за другой, использовав для него символ $\vdash$.

Естественно ввести степень отношения $\overset{n}{\vdash}$, транзитивное замыкание $\overset{+}{\vdash}$ и рефлексивно-транзитивное замыкание $\overset{*}{\vdash}$ этого отношения на множестве конфигураций. При необходимости явно показать, движения какого pda рассматриваются, под соответствующими значками указывается имя автомата, например: $\underset{M}{\vdash}$, $\overset{n}{\underset{M}{\vdash}}$, $\overset{+}{\underset{M}{\vdash}}$ или $\overset{*}{\underset{M}{\vdash}}$. Напомним, что содержательно эти значки обозначают переход от конфигурации, стоящей слева от значка, к конфигурации, стоящей справа от значка, соответственно за одно движение, за $n$ движений, за положительное число движений, за произвольное, может быть нулевое, число движений.

***<u>Определение</u>* <u>5.3.</u>** *Язык, принимаемый магазинным автоматом* $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ *при конечном состоянии*, определим как множество $T(M) = \{w \in \Sigma^* \mid (q_0, w, Z_0) \overset{*}{\vdash} (q, \varepsilon, \alpha), q \in F\}$.

***<u>Определение</u>* <u>5.4.</u>** *Язык, принимаемый магазинным автоматом* $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ *при пустом магазине*, определим как множество $N(M) = \{w \in \Sigma^* \mid (q_0, w, Z_0) \overset{*}{\vdash} (q, \varepsilon, \varepsilon), q \in Q\}$. В этом случае неважно, каким является множество конечных состояний, и часто оно просто считается пустым.

**Пример 5.1.** Пусть pda $M = (\{q_1, q_2\}, \{0, 1\}, \{R, B, G\}, \delta, q_1, R, \varnothing)$, где

1) $\delta(q_1, 0, R) = \{(q_1, BR)\}$,
2) $\delta(q_1, 1, R) = \{(q_1, GR)\}$,
3) $\delta(q_1, 0, B) = \{(q_1, BB), (q_2, \varepsilon)\}$,
4) $\delta(q_1, 0, G) = \{(q_1, BG)\}$,
5) $\delta(q_1, 1, B) = \{(q_1, GB)\}$,
6) $\delta(q_1, 1, G) = \{(q_1, GG), (q_2, \varepsilon)\}$,
7) $\delta(q_2, 0, B) = \{(q_2, \varepsilon)\}$,
8) $\delta(q_2, 1, G) = \{(q_2, \varepsilon)\}$,
9) $\delta(q_1, \varepsilon, R) = \{(q_2, \varepsilon)\}$,
10) $\delta(q_2, \varepsilon, R) = \{(q_2, \varepsilon)\}$;

$M$ — недетерминированный магазинный автомат, принимающий язык $N(M) = \{ww^R \mid w \in \{0, 1\}^*\}$, где $w^R$ обозначает инвертированную цепочку $w$, при пустом магазине.

Правила 1–6 позволяют pda $M$ запомнить входной символ в магазинной памяти. При этом входной символ 0 отображается в магазине посредством символа $B$, а символ 1 отображается в магазине посредством символа $G$.

Правила 3 и 6 предоставляют автомату выбор движений между запоминанием очередного входного символа в магазине (в предположении, что середина входной цепочки не достигнута) и переходом в режим сопоставления текущего входного символа с символом на вершине магазина и стиранием последнего в случае их соответствия (в предположении, что достигнута середина входной цепочки).

Если на входе находится цепочка вида $ww^R$, то существует такая последовательность движений, которая опустошает магазин к моменту достижения конца цепочки. Например, если на вход поступает цепочка 0110, то автомат имеет выбор движений, представленный на рис. 5.1, где приведено дерево возможных переходов между конфигурациями. Существующая среди них следующая последовательность движений:

$$(q_1, 0110, R) \vdash (q_1, 110, BR) \vdash (q_1, 10, GBR) \vdash (q_2, 0, BR) \vdash (q_2, \varepsilon, R) \vdash (q_2, \varepsilon, \varepsilon)$$

дает основание заключить, что входная цепочка 0110 принимается.

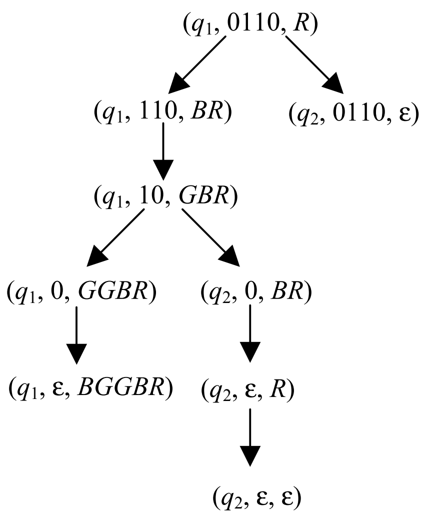


Рис. 5.1.

**Пример 5.2.** Пусть pda $M = (\{q_1, q_2\}, \{0, 1, c\}, \{R, B, G\}, \delta, q_1, R, \varnothing)$, где

1) $\delta(q_1, 0, R) = \{(q_1, BR)\}$,
2) $\delta(q_1, 0, B) = \{(q_1, BB)\}$,
3) $\delta(q_1, 0, G) = \{(q_1, BG)\}$,
4) $\delta(q_1, c, R) = \{(q_2, R)\}$,
5) $\delta(q_1, c, B) = \{(q_2, B)\}$,
6) $\delta(q_1, c, G) = \{(q_2, G)\}$,
7) $\delta(q_2, 0, B) = \{(q_2, \varepsilon)\}$,
8) $\delta(q_2, \varepsilon, R) = \{(q_2, \varepsilon)\}$,
9) $\delta(q_2, 1, G) = \{(q_2, \varepsilon)\}$,
10) $\delta(q_1, 1, R) = \{(q_1, GR)\}$,
11) $\delta(q_1, 1, B) = \{(q_1, GB)\}$,
12) $\delta(q_1, 1, G) = \{(q_1, GG)\}$;

Легко сообразить, что $N(M) = \{wcw^R \mid w \in \{0,1\}^*\}$, где $w^R$ обозначает инвертированную цепочку $w$.

Рассмотрим движения pda $M$ при наличии цепочки $01c10$ на его входе:

$$(q_1, 01c10, R) \vdash_{M} (q_1, 1c10, BR) \vdash_{M} (q_1, c10, GBR) \vdash_{M}$$
$$(q_2, 10, GBR) \vdash_{M} (q_2, 0, BR) \vdash_{M} (q_2, \varepsilon, R) \vdash_{M} (q_2, \varepsilon, \varepsilon).$$

Следовательно, цепочка $01c10$ принимается при пустом магазине.

Заметим, что равенство $\delta(q_2, \varepsilon, R) = \{(q_2, \varepsilon)\}$ означает движение, не зависящее от входного символа, каким бы он ни был, — в любом случае происходит стирание верхнего символа $R$ магазина без продвижения по входной цепочке и переход в состояние $q_2$.

Магазинный автомат из примера 5.2 является детерминированным в том смысле, что из любой конфигурации возможно не более одного движения.

***Определение* 5.5.** Магазинный автомат $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ является детерминированным, если

1) для любых $q \in Q$, $Z \in \Gamma$ и $a \in (\Sigma \cup \{\varepsilon\})$ значение $\#\delta(q, a, Z) \le 1$;

2) для любых $q \in Q$ и $Z \in \Gamma$ всякий раз, как $\delta(q, \varepsilon, Z) \ne \varnothing$, значение $\delta(q, a, Z) = \varnothing$ для всех $a \in \Sigma$.

Условие 1 означает, что если движение определено, то оно единственно. Условие 2 предотвращает выбор между $\varepsilon$-движением и движением, использующим входной символ.

Для конечных автоматов детерминированная и недетерминированная модели эквивалентны по отношению к применяемым языкам (см. теорему 3.4). Далее мы увидим, что это не так для МП-автоматов. Действительно, язык, состоящий из цепочек вида $ww^R$, принимается недетерминированным pda, но не существует никакого детерминированного pda, который принимал бы такой язык.

## § 5.3. Недетерминированные магазинные автоматы и контекстно-свободные языки

Теперь докажем фундаментальный результат: класс языков, принимаемых недетерминированными МП-автоматами, есть в точности класс контекстно-свободных языков. Для этого сначала покажем, что языки, принимаемые недетерминированными МП-автоматами при конечном состоянии, являются в точности языками, принимаемыми недетерминированными МП-автоматами при пустом магазине, а затем, что языки, принимаемые при пустом магазине, совпадают с классом КС-языков.

**Теорема 5.1.** *Язык* $L = N(M_1)$ *для некоторого недетерминированного магазинного автомата* $M_1$ *тогда и только тогда, когда* $L = T(M_2)$ *для некоторого магазинного автомата* $M_2$.

Доказательство.

I. *Необходимость*. Пусть $M_1 = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, \varnothing)$ — недетерминированный pda, такой, что $L = N(M_1)$. Положим

$$M_2 = (Q \cup \{q_0', q_f\}, \Sigma, \Gamma \cup \{X\}, \delta', q_0', X, \{q_f\}),$$

где $\delta'$ определяется следующим образом:

1) $\delta'(q_0', \varepsilon, X) = \{(q_0, Z_0X)\}$;

2) $\delta'(q, a, Z) = \delta(q, a, Z)$ для всех $q \in Q$, $a \in (\Sigma \cup \{\varepsilon\})$ и $Z \in \Gamma$;

3) $\delta'(q, \varepsilon, X) = \{(q_f, \varepsilon)\}$ для всех $q \in Q$.

Правило 1 воспроизводит начальную конфигурацию автомата $M_1$. При этом начальный символ магазина $X$ остается в качестве своеобразного маркера дна магазина до тех пор, пока он не будет удален последним движением автомата $M_2$. Выше этого маркера при последующих движениях, повторяющих движения pda $M_1$, всегда будут находиться только магазинные символы автомата $M_1$. Это обеспечивается правилом 2. Когда же будет воспроизведено последнее движение pda $M_1$, опустошающее его магазин, на вершине магазина автомата $M_2$ появится маркер дна. В этот момент автомат $M_2$ посредством правила 3 переводится в конечное состояние, тем самым сигнализируя прием входной цепочки.

Заметим, что движение по правилу 3 стирает символ магазина $X$, но могло бы записать вместо него любую магазинную цепочку.

I.1. Докажем, сначала, что если $x \in N(M_1)$, то $x \in T(M_2)$.

Пусть $x \in N(M_1)$, т.е. $(q_0, x, Z_0) \vdash^*_{M_1} (q, \varepsilon, \varepsilon)$. Посмотрим, какие движения будет совершать автомат $M_2$, имея на входе цепочку символов $x$. Согласно правилу 1 $(q_0', x, X) \vdash_{M_2} (q_0, x, Z_0X)$. Далее согласно правилу 2 автомат $M_2$, повторяя движения автомата $M_1$, осуществляет переход $(q_0, x, Z_0X) \vdash^*_{M_2} (q, \varepsilon, X)$. Наконец, по правилу 3 имеем последнее движение: $(q, \varepsilon, X) \vdash_{M_2} (q_f, \varepsilon, \varepsilon)$. Таким образом, $x \in T(M_2)$.

I.2. Теперь докажем, что если $x \in T(M_2)$ , то $x \in N(M_1)$.

Пусть $x \in T(M_2)$, т.е. $(q_0', x, X) \vdash^*_{M_2} (q_f, \varepsilon, \gamma)$ для некоторой цепочки $\gamma \in (\Gamma \cup \{X\})^*$, причем по построению автомата $M_2$ $\gamma = \varepsilon$. Выделив здесь первый шаг явным образом, имеем $(q_0', x, X) \vdash_{M_2} (q_0, x, Z_0X) \vdash^*_{M_2} (q_f, \varepsilon, \varepsilon)$. Заключительная конфигурация $(q_f, \varepsilon, \varepsilon)$ достижима лишь посредством $\varepsilon$-движения благодаря правилу 3 в случае, когда на вершине магазина находится $X$. Следовательно, предпоследняя конфигурация должна иметь вид $(q_0, \varepsilon, X)$. Итак, все движения выстраиваются в следующую последовательность конфигураций:

$$(q_0', x, X) \vdash_{M_2} (q_0, x, Z_0X) \vdash_{M_2}^{*} (q, \varepsilon, X) \vdash_{M_2} (q_f, \varepsilon, \varepsilon).$$

Но на главном участке $(q_0, x, Z_0X) \vdash_{M_2}^{*} (q, \varepsilon, X)$ автомат $M_2$ просто повторяет движения автомата $M_1$. Поэтому $(q_0, x, Z_0) \vdash_{M_1}^{*} (q, \varepsilon, \varepsilon)$ и $x \in N(M_1)$.

Из рассуждений, проведенных в пп. I.1 и I.2 следует, что $T(M_2) = N(M_1)$.

II. *Достаточность*. Пусть $M_2 = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ — недетерминированный pda, такой, что $L = T(M_2)$. Положим

$$M_1 = (Q \cup \{q_0', q_e\}, \Sigma, \Gamma \cup \{X\}, \delta', q_0', X, \varnothing),$$

где $\delta'$ определяется следующим образом:

1) $\delta'(q_0', \varepsilon, X) = \{(q_0, Z_0X)\}$;

2) $\delta'(q, a, Z) = \delta(q, a, Z)$ для всех $q \in Q$, $a \in (\Sigma \cup \{\varepsilon\})$ и $Z \in \Gamma$;

3) $\delta'(q, \varepsilon, Z) = \{(q_e, \varepsilon)\}$ для всех $q \in F$ и $Z \in \Gamma \cup \{X\}$;

4) $\delta'(q_e, \varepsilon, Z) = \{(q_e, \varepsilon)\}$ для всех $Z \in \Gamma \cup \{X\}$.

Правило 1 воспроизводит начальную конфигурацию автомата $M_2$. При этом начальный символ магазина $X$ остается в качестве своеобразного маркера дна магазина до тех пор, пока он не будет удален последним движением автомата $M_1$. Выше этого маркера при последующих движениях, повторяющих движения автомата $M_2$, всегда будут находиться только магазинные символы автомата $M_2$. Это обеспечивается правилом 2. Если когда-либо pda $M_2$ входит в конечное состояние, то правила 3 и 4 позволяют pda $M_1$ выбирать переход в состояние $q_e$ и, не продвигаясь по входу, очищать магазин, тем самым принимая вход или продолжая моделировать движения автомата $M_2$.

Следует заметить, что автомат $M_2$ может на неправильном входе опустошить свой магазин, не заходя в конечное состояние и, следовательно, не принимая входную цепочку. Чтобы неправильные цепочки не принимались pda $M_1$, и введено его собственное особое дно в виде символа $X$, чтобы он, воспроизводя эти же движения, тоже не опустошил свой магазин и таким образом принял бы неправильную цепочку.

Остается убедиться в том, что $N(M_1) = T(M_2)$.

II.1. Докажем, сначала, что если $x \in N(M_1)$, то $x \in T(M_2)$.

Если $x \in N(M_1)$, то по определению $(q_0', x, X) \vdash_{M_1}^{*} (q, \varepsilon, \varepsilon)$ при некотором $q \in Q \cup \{q_0', q_e\}$. Рассмотрим подробнее эти движения. Согласно правилу 1 имеем $(q_0', x, X) \vdash_{M_1} (q_0, x, Z_0X)$. Далее $(q_0, x, Z_0X) \vdash_{M_1}^{*} (q, \varepsilon, \varepsilon)$ при некотором $q \in Q \cup \{q_0', q_e\}$.

Удалить $X$ из магазина согласно построению автомата $M_1$ возможно только посредством правил 3 или 4, причем правило 4 применимо только после того, как применено правило 3. В свою очередь, правило 3 применимо только в том случае, если автомат $M_1$ достиг одного из своих конечных состояний. С этого момента происходят только ε-движения, стирающие содержимое магазина. С учетом этого имеем

$$(q_0', x, X) \vdash_{M_1} (q_0, x, Z_0X) \vdash_{M_1}^{*} (q, \varepsilon, \gamma X) \vdash_{M_1} (q_e, \varepsilon, \gamma') \vdash_{M_1}^{*} (q_e, \varepsilon, \varepsilon).$$

Здесь $q \in F$, $\gamma \in \Gamma^*$, $\gamma' \in (\Gamma \cup \{X\})^*$, причем если $\gamma' = \varepsilon$, то перехода $(q_e, \varepsilon, \gamma') \vdash_{M_1}^{*}$ $(q_e, \varepsilon, \varepsilon)$ фактически нет.

Но мы знаем, что согласно правилу 2 автомат $M_2$ на главном участке просто повторяет движения автомата $M_1$: $(q_0, x, Z_0X) \vdash_{M_1}^{*} (q, \varepsilon, \gamma X)$, т.е. имеет место переход $(q_0, x, Z_0X) \vdash_{M_2}^{*} (q, \varepsilon, \gamma X)$, где $q \in F$. Следовательно, $x \in T(M_2)$.

II.2. Теперь докажем, что если $x \in T(M_2)$, то $x \in N(M_1)$.

Пусть $x \in T(M_2)$, т.е. $(q_0, x, Z_0) \vdash_{M_2}^{*} (q, \varepsilon, \gamma)$, где $q \in F$, $\gamma \in \Gamma^*$. Посмотрим, как будет действовать автомат $M_1$ на том же входе. Согласно правилу 1 имеем $(q_0', x, X) \vdash_{M_1}$ $(q_0, x, Z_0X)$. Затем согласно правилу 2 воспроизводятся указанные ранее движения автомата $M_2$, т.е. имеем $(q_0', x, X) \vdash_{M_1} (q_0, x, Z_0X) \vdash_{M_1}^{*} (q, \varepsilon, \gamma X)$, где $q \in F$, $\gamma \in \Gamma^*$. Далее согласно правилам 3 и 4 имеем $(q, \varepsilon, \gamma X) \vdash_{M_1} (q_e, \varepsilon, \gamma') \vdash_{M_1}^{*} (q_e, \varepsilon, \varepsilon)$, где $\gamma'(\Gamma \cup \{X\})^*$. Если $\gamma' = \varepsilon$, то последнего перехода фактически нет.

Итак, существует последовательность конфигураций:

$$(q_0', x, X) \vdash_{M_1} (q_0, x, Z_0X) \vdash_{M_1}^{*} (q, \varepsilon, \gamma X) \vdash_{M_1} (q_e, \varepsilon, \gamma') \vdash_{M_1}^{*} (q_e, \varepsilon, \varepsilon),$$

означающая, что $x \in N(M_1)$.

Из рассуждений. проведенных в пп. II.1 и II.2 следует, что $N(M_1) = T(M_2)$.

Заключение теоремы следует из доказанных утверждений I и II.

**Теорема 5.2.** *Если $L$ — контекстно-свободный язык, то существует недетерминированный магазинный автомат $M$, такой, что $L = N(M)$.*

Доказательство. Пусть $G = (V_N, V_T, P, S)$ — контекстно-свободная грамматика в нормальной форме Грейбах, и $L = L(G)$. Мы предполагаем, что $\varepsilon \notin L(G)$.

Пусть $S \underset{\text{lm}}{\overset{*}{\Rightarrow}} x\alpha$, где $x \in V_T^*$, причем $\alpha = A\beta$, $A \in V_N$, $\beta \in V_N^*$, либо $\alpha = \varepsilon$. Цепочка $x$ называется *закрытой*, а $\alpha$ — *открытой частью* сентенциальной формы $x\alpha$.

Построим недетерминированный pda $M$ таким образом, чтобы в его магазине воспроизводились последовательности сентенциальных форм, составляющих левосторонние выводы в грамматике $G$, вернее, открытые части таких форм. Для этого положим $M = (\{q\}, V_T, V_N, \delta, q, S, \varnothing)$, где $(q, \gamma) \in \delta(q, a, A)$, если существует правило $A \to a\gamma \in P$. Очевидно, что в магазине pda $M$ всегда находится нетерминальная цепочка, причем на вершине — символ, подлежащий замене на текущем шаге левостороннего вывода.

Чтобы показать, что $L(G) = N(M)$, достаточно доказать, что $A \underset{\mathrm{lm}}{\overset{*}{\Rightarrow}} x\alpha$ тогда и только тогда, когда $(q, x, A) \overset{*}{\underset{M}{\vdash}} (q, \varepsilon, \alpha)$ для любых $x \in V_T^*$, $A \in V_N$, $\alpha \in V_N^*$. Утверждение теоремы последует как частный случай этого вспомогательного утверждения при $A = S$, $\alpha = \varepsilon$: $S \underset{G}{\overset{*}{\Rightarrow}} x$ тогда и только тогда, когда $(q, x, S) \underset{M}{\vdash} (q, \varepsilon, \varepsilon)$.

Заметим, что pda $M$ не совершает ε-движений.

I. Индукцией по длине вывода $l$ покажем, что если $A \underset{G}{\overset{l}{\Rightarrow}} x\alpha$, где $x \in V_T^*$, $\alpha \in V_N^*$, то $(q, x, A) \overset{*}{\underset{M}{\vdash}} (q, \varepsilon, \alpha)$.

База. Пусть $l = 1$. Имеем $A \underset{G}{\Rightarrow} x\alpha$. Это значит, что существует правило $A \to x\alpha \in P$, где $x \in V_T$, $\alpha \in V_N^*$. Одновременно по построению автомата $M$ мы имеем $(q, \alpha) \in \delta(q, x, A)$, а тогда $(q, x, A) \underset{M}{\vdash} (q, \varepsilon, \alpha)$.

Индукционная гипотеза. Предположим, что для всех $l \le n$ $(n \ge 1)$, если $A \underset{G}{\overset{l}{\Rightarrow}} x\alpha$, то $(q, x, A) \overset{*}{\underset{M}{\vdash}} (q, \varepsilon, \alpha)$.

Индукционный переход. Докажем это утверждение для $l = n + 1$. Пусть $A \underset{G}{\overset{n}{\Rightarrow}} yB\gamma \underset{G}{\Rightarrow} ya\beta\gamma = x\alpha$ — левосторонний вывод длиной $n + 1$. В нем $x = ya$, $\alpha = \beta\gamma$ и $B \to a\beta \in P$. По построению автомата $M$ существует $(q, \beta) \in \delta(q, a, B)$. В соответствии с индукционной гипотезой и с учетом этого последнего обстоятельства можем написать: $(q, x, A) = (q, ya, A) \overset{*}{\underset{M}{\vdash}} (q, a, B\gamma) \underset{M}{\vdash} (q, \varepsilon, \beta\gamma) = (q, \varepsilon, \alpha)$.

II. Теперь индукцией по числу $l$ движений автомата $M$ покажем, что если $(q, x, A) \overset{l}{\underset{M}{\vdash}} (q, \varepsilon, \alpha)$, где $x \in V_T^*$, $\alpha \in V_N^*$, то $A \underset{G}{\overset{*}{\Rightarrow}} x\alpha$.

База. Пусть $l = 1$. Имеем $(q, x, A) \underset{M}{\vdash} (q, \varepsilon, \alpha)$, где $x \in V_T^*$, $\alpha \in V_N^*$. Поскольку pda $M$ не совершает ε-движений, то это движение имеет место благодаря тому, что $(x, \alpha) \in \delta(q, x, A)$, где $x \in V_T$. Но это обусловлено тем, что существует правило $A \to x\alpha \in P$, с помощью которого получаем требуемый вывод $A \underset{G}{\Rightarrow} x\alpha$.

Индукционная гипотеза. Предположим, что утверждение выполняется для всех $l \le n$ $(n \ge 1)$.

Индукционный переход. Докажем это утверждение для $l = n + 1$. Пусть $(q, x, A) = (q, ya, A) \overset{n}{\underset{M}{\vdash}} (q, a, B\gamma) \underset{M}{\vdash} (q, \varepsilon, \beta\gamma) = (q, \varepsilon, \alpha)$, где $x = ya$, $\alpha = \beta\gamma$. Последнее из этих движений подразумевает, что $(q, \beta) \in \delta(q, a, B)$, где $a \in V_T$, а это обусловлено существованием правила $B \to a\beta \in P$. Учитывая следствие индукционной гипотезы из $(q, ya, A) \overset{n}{\underset{M}{\vdash}} (q, a, B\gamma)$ и упомянутое правило, мы можем написать:

$$A \underset{G}{\overset{*}{\Rightarrow}} yB\gamma \underset{G}{\Rightarrow} ya\beta z\gamma = x\alpha.$$

Из рассуждений, проведенных в пп. I и II, при $A = S$ и $\alpha = \varepsilon$ получаем утверждение теоремы. Что и требовалось доказать.

**Теорема 5.3.** *Если $M$ — недетерминированный магазинный автомат, и $L = N(M)$, то $L$ — контекстно-свободный язык.*

Доказательство. Пусть $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, \varnothing)$ — недетерминированный pda, такой, что $L = N(M)$.

Построим cfg $G$, левосторонние выводы в которой моделируют движения pda $M$. В частности, нетерминалы, которые появляются в сентенциальных формах на любом шаге левостороннего вывода в грамматике $G$, соответствуют символам в магазине автомата $M$ в момент, когда он уже просмотрел такую же часть входной цепочки, какую грамматика уже породила.

Положим $G = (V_N, V_T, P, S)$, где $V_N = \{S\} \cup \{[qAp] \mid q, p \in Q;\ A \in \Gamma\}$; $V_T = \Sigma$;

$$P = \{(1)\ S \to [q_0 Z_0 q] \mid q \in Q\} \cup$$
$$\{(2)\ [qAp] \to a\,[q_1B_1q_2]\,[q_2B_2q_3]...[q_mB_mq_{m+1}] \mid q,\ qz_i \in Q\ (i = 1, 2, \dots, m+1);$$
$$p = q_{m+1};\ a \in \Sigma \cup \{\varepsilon\};\ A, B_j \in \Gamma\ (j = 1, 2, \dots, m);\ (q_1, B_1B_2...B_m) \in \delta(q, a, A)\}.$$

Отметим, что если $(q_1, \varepsilon) \in \delta(q, a, A)$, то $m = 0$, $q_1 = p$ и правило вида 2 имеет вид $[qAp] \to a$.

Чтобы показать, что $L(G) = N(M)$, мы сначала докажем вспомогательное утверждение: $[qAp] \overset{*}{\underset{G}{\Rightarrow}} x$ тогда и только тогда, когда $(q, x, A) \overset{*}{\underset{M}{\vdash}} (p, \varepsilon, \varepsilon)$.

I. Индукцией по числу $l$ движений pda $M$ покажем, что если $(q, x, A) \overset{*}{\underset{M}{\vdash}} (p, \varepsilon, \varepsilon)$, то $[qAp] \overset{*}{\underset{G}{\Rightarrow}} x$.

База. Пусть $l = 1$. Имеем $(q, x, A) \underset{M}{\vdash} (p, \varepsilon, \varepsilon)$. Это движение обусловлено тем, что $(p, \varepsilon) \in \delta(q, x, A)$, где $x \in \Sigma \cup \{\varepsilon\}$. Соответственно по способу построения грамматики $G$ существует правило $[qAp] \to x$, при помощи которого $[qAp] \underset{G}{\Rightarrow} x$.

Индукционная гипотеза. Предположим, что утверждение выполняется для всех значений $l \le n$ $(n \ge 1)$.

Индукционный переход. Докажем, что утверждение выполняется для $l = n + 1$. Пусть $(q, x, A) = (q, ax_1x_2... x_n, A) \underset{M}{\vdash} (q_1, x_1x_2... x_n, B_1B_2...B_m) \overset{n}{\underset{M}{\vdash}} (p, \varepsilon, \varepsilon)$. Здесь $x = ax_1x_2 \dots x_n$, $a \in \Sigma \cup \{\varepsilon\}$, $x_i \in \Sigma^*$, причем $(q_i, x_i, B_i) \overset{l_i}{\underset{M}{\vdash}} (q_{i+1}, \varepsilon, \varepsilon)$, $i = 1, 2, \dots, m$, $q_{m+1} = p$.

Поясним, что в этой последовательности переходов явно показано первое движение, которое может использовать входной символ $a$, с которого начинается входная цепочка $x$, либо оно может быть $\varepsilon$-движением. Каждая из подцепочек $x_i$ есть та часть цепочки $x$, которую автомат просматривает с момента, когда он оказывается в состоянии $q_i$ с символом $B_i$ на вершине магазина, до того момента, когда вершина магазина впервые опустится на один символ ниже этой позиции $B_i$.

Очевидно, что первое движение обусловлено тем, что $(q_1, B_1B_2 \dots B_m) \in \delta(q, a, A)$, и, как следствие этого, существует правило грамматики вида

$$[qAp] \to a\,[q_1B_1q_2][q_2B_2q_3]...[q_mB_mq_{m+1}]$$

с нетерминалами, связанными с состояниями и магазинными символами, участвующими в указанных конфигурациях. С другой стороны, применение индукционной гипотезы к переходу $(q_i, x_i, B_i) \vdash_M^{l_i} (q_{i+1}, \varepsilon, \varepsilon)$, где $l_i \le n$, дает основание заключить, что $[q_iB_iq_{i+1}] \overset{*}{\underset{G}{\Rightarrow}} x_i$, $i = 1, 2, \ldots, m$, $q_{m+1} = p$.

С помощью упомянутых правила и частичных выводов можно выстроить требуемый вывод:

$[qAp] \underset{G}{\Rightarrow} a[q_1B_1q_2][q_2B_2q_3]\ldots[q_mB_mq_{m+1}] \overset{*}{\underset{G}{\Rightarrow}} ax_1[q_2B_2q_3]\ldots[q_mB_mq_{m+1}] \overset{*}{\underset{G}{\Rightarrow}} \ldots$
$\overset{*}{\underset{G}{\Rightarrow}} ax_1x_2\ldots x_n = x.$

II. Индукцией по длине вывода $l$ покажем, что если $[qAp] \overset{l}{\underset{G}{\Rightarrow}} x$, то $(q, x, A) \vdash_M^{*} (p, \varepsilon, \varepsilon)$.

База. Пусть $l = 1$ . Имеем $[qAp] \underset{G}{\Rightarrow} x$. Очевидно, что на этом единственном шаге применено правило $[qAp] \to x$. По способу построения cfg $G$ $x \in \Sigma \cup \{\varepsilon\}$. Такое правило существует, если только $(p, \varepsilon) \in \delta(q, x, A)$. А тогда $(q, x, A) \vdash_M (p, \varepsilon, \varepsilon)$.

Индукционная гипотеза. Предположим, что утверждение выполняется для всех значений $l \le n$ $(n \ge 1)$.

Индукционный переход. Докажем, что утверждение выполняется для $l = n + 1$. Пусть $[qAp] \underset{G}{\Rightarrow} a[q_1B_1q_2][q_2B_2q_3]\ldots[q_mB_mq_{m+1}] \overset{n}{\underset{G}{\Rightarrow}} ax_1x_2\ldots x_n = x$. Очевидно, что существует правило $[qAp] \to a[q_1B_1q_2][q_2B_2q_3]\ldots[q_mB_mq_{m+1}] \in P$, использованное на первом шаге вывода, в котором $q_{m+1} = p$, и частичные выводы $[q_iB_iq_{i+1}] \overset{l_i}{\underset{G}{\Rightarrow}} x_i$, где $l_i \le n$, $i = 1, 2, \ldots, m$. Из существования такого правила следует, что $(q_1, B_1B_2\ldots B_m) \in \delta(q, a, A)$. Из частичных выводов в соответствии с индукционной гипотезой вытекает, что $(q_i, x_i, B_i) \vdash_M^{l_i} (q_{i+1}, \varepsilon, \varepsilon)$, $i = 1, 2, \ldots, m$, $q_{m+1} = p$. Следовательно, существует последовательность конфигураций:

$$(q, x, A) = (q, ax_1x_2\ldots x_n, A) \vdash_M (q_1, x_1x_2\ldots x_n, B_1B_2\ldots B_m) \vdash_M^{*}$$
$$\vdash_M^{*} (q_2, x_2\ldots x_n, B_2\ldots B_m) \vdash_M^{*} \ldots \vdash_M^{*} (p, \varepsilon, \varepsilon).$$

Итак, из рассуждений, проведенных в пп. I и II, следует справедливость вспомогательного утверждения: $[qAp] \overset{*}{\underset{G}{\Rightarrow}} x$ тогда и только тогда, когда $(q, x, A) \vdash_M^{*} (p, \varepsilon, \varepsilon)$. В частности, при $q = q_0$ и $A = Z_0$ $[q_0Z_0p] \overset{*}{\underset{G}{\Rightarrow}} x$ тогда и только тогда, когда $(q_0, x, Z_0) \vdash_M^{*} (p, \varepsilon, \varepsilon)$.

Остается вспомнить, что существуют правила вида $S \to [q_0Z_0q]$ для любых $q \in Q$, в частности для $q = p$; тогда, пристраивая новое начало к выводу, получаем справедливость следующего утверждения: $S \underset{G}{\Rightarrow} [q_0Z_0p] \overset{*}{\underset{G}{\Rightarrow}} x$ тогда и только тогда, когда $(q_0, x, Z_0) \vdash_M^{*} (p, \varepsilon, \varepsilon)$, или, что то же самое, $L(G) = N(M)$. Что и требовалось доказать.

Теоремы 5.1, 5.2 и 5.3 можно подытожить: следующие три утверждения являются эквивалентными:

1) $L$ — cfl;
2) $L = N(M_1)$ для некоторого pda $M_1$;
3) $L = T(M_2)$ для некоторого pda $M_2$.

**Пример 5.3.** Пусть pda $M = (\{q_0, q_1\}, \{0, 1\}, \{X, Z_0\}, \delta, q_0, Z_0, \varnothing)$,

где $\delta = \{$ (1) $\delta(q_0, 0, Z_0) = \{(q_0, XZ_0)\}$, (4) $\delta(q_1, 1, X) = \{(q_1, \varepsilon)\}$,

(2) $\delta(q_0, 0, X) = \{(q_0, XX)\}$, (5) $\delta(q_1, \varepsilon, X) = \{(q_1, \varepsilon)\}$,

(3) $\delta(q_0, 1, X) = \{(q_1, \varepsilon)\}$, (6) $\delta(q_1, \varepsilon, Z_0) = \{(q_1, \varepsilon)\}\}$.

Нетрудно сообразить, что $N(M) = \{0^n 1^m \mid n \geq m\}$.

Построим cfg $G = (V_N, V_T, P, S)$, порождающую язык $N(M)$, положив

$V_N = \{S, [q_0 Z_0 q_0], [q_0 Z_0 q_1], [q_1 Z_0 q_0], [q_1 Z_0 q_1], [q_0 X q_0], [q_0 X q_1], [q_1 X q_0], [q_1 X q_1]\}$,

$V_T = \{0, 1\}$.

Для экономии времени при построении правил грамматики следует учесть, что некоторые нетерминалы не участвуют ни в каких выводах, начинающихся с символа $S$. Поэтому мы должны начать с построения правил для $S$, затем построить правила для тех нетерминалов, которые встречаются в правых частях уже построенных правил.

Правилами для $S$ являются: $S \to [q_0 Z_0 q_0]$ и $S \to [q_0 Z_0 q_1]$.

Затем, исходя из правила 1 в определении $\delta$, мы добавим правила для нетерминала $[q_0 Z_0 q_0]$:

$$[q_0 Z_0 q_0] \to 0[q_0 X q_0][q_0 Z_0 q_0],$$

$$[q_0 Z_0 q_0] \to 0[q_0 X q_1][q_1 Z_0 q_0].$$

Так как в правых частях построенных правил появился нетерминал $[q_0 Z_0 q_1]$, мы должны построить правила для него, опять исходя из того же правила 1:

$$[q_0 Z_0 q_1] \to 0[q_0 X q_0][q_0 Z_0 q_1],$$

$$[q_0 Z_0 q_1] \to 0[q_0 X q_1][q_1 Z_0 q_1].$$

Аналогично, исходя из правила 2 в определении $\delta$, получаем следующие правила грамматики:

$$[q_0 X q_0] \to 0[q_0 X q_0][q_0 X q_0],$$

$$[q_0 X q_0] \to 0[q_0 X q_1][q_1 X q_0],$$

$$[q_0 X q_1] \to 0[q_0 X q_0][q_0 X q_1],$$

$$[q_0 X q_1] \to 0[q_0 X q_1][q_1 X q_1].$$

Исходя из правил 3–6 в определении $\delta$, получаем следующие правила грамматики:

$$[q_0 X q_1] \to 1,\ [q_1 X q_1] \to 1,\ [q_1 X q_1] \to \varepsilon,\ [q_1 Z_0 q_1] \to \varepsilon.$$

Заметим, что нет никаких правил для нетерминалов $[q_1 Z_0 q_0]$ и $[q_1 X q_0]$. Соответственно правила, в правых частях которых встречаются эти два нетерминала, бесполезны и могут быть отброшены. Кроме того, легко заметить, что из нетерминалов $[q_0 Z_0 q_0]$ и $[q_0 X q_0]$ не выводится ни одной терминальной цепочки. Поэтому правила с их участием могут быть отброшены.

Окончательно получаем следующее множество правил приведенной грамматики:

(1) $S \to [q_0 Z_0 q_1]$, (4) $[q_0 X q_1] \to 1$,

(2) $[q_0 Z_0 q_1] \to 0[q_0 X q_1][q_1 Z_0 q_1]$, (5) $[q_1 X q_1] \to 1$,

(3) $[q_0 X q_1] \to 0[q_0 X q_1][q_1 X q_1]$, (6) $[q_1 X q_1] \to \varepsilon$, (7) $[q_1 Z_0 q_1] \to \varepsilon$.