Часть II: ТРАНСЛЯЦИИ И СИНТАКСИЧЕСКИЕ МЕТОДЫ ИХ РЕАЛИЗАЦИИ

Глава 1

ТРАНСЛЯЦИИ, ИХ ПРЕДСТАВЛЕНИЕ И РЕАЛИЗАЦИЯ

§ 1.1. Трансляции и трансляторы

Определение **1.1.** *Трансляцией* из языка $L_1 \subseteq \Sigma^*$ в язык $L_2 \subseteq \Delta^*$ называется отношение $\tau \subseteq L_1 \times L_2$. Здесь Σ — *входной алфавит*, L_1 — *входной язык*, Δ — *выходной алфавит*, L_2 — *выходной язык*.

Другими словами, трансляция есть некоторое множество пар предложений (x, y), где $x \in L_1$ — *входное*, а $y \in L_2$ — *выходное предложение*. Хотя в общем случае в трансляции τ одному входному предложению x может соответствовать несколько выходных предложений y, по отношению к языкам программирования трансляция всегда является функцией, т.е. для каждого входа существует не более одного выхода.

Существует бесконечно много примеров трансляций, но самым элементарным, вероятно, является тот, который может быть задан *гомоморфизмом*, т.е. отображением h из Σ в Δ^*.

Пример 1.1. Предположим, что мы хотим закодировать некоторый текст с помощью азбуки Морзе. Как известно, в коде Морзе каждая буква представляется как некоторая последовательность из точек и тире. Эти последовательности, называемые *посылками*, имеют разную длину. Для отделения одной посылки от другой используется *пауза*. Очевидно, что трансляцию предложений, например на русском языке, в код Морзе можно реализовать с помощью гомоморфизма, задаваемого следующим образом:

Буква	а	б	в	…	я
Посылка	•—	—•••	•——	…	•—•—

Для простоты предполагаем, что паузы представлены пробелами, завершающими каждую посылку. Тогда, скажем, слово “абба” с помощью замены букв на посылки даст результат:

•— —••• —••• •—

Для любой входной цепочки $x = a_1 a_2 \ldots a_n$, $a_i \in \Sigma$, $i = 1, 2, \ldots, n$, гомоморфизм h позволяет найти соответствующую выходную цепочку y с помощью посимвольной подстановки: $y = h(a_1)h(a_2)\ldots h(a_n)$.

Область определения гомоморфизма можно расширить до Σ^* следующим образом: $h(ax) = h(a)h(x)$, $h(\varepsilon) = \varepsilon$. Здесь $a \in \Sigma$, $x \in \Sigma^*$.

Гомоморфизм h определяет трансляцию $\tau(h) = \{(x,\ h(x)) \mid x \in \Sigma^*\}$.

Устройство, которое по заданной цепочке $x \in \Sigma^*$ находит соответствующую цепочку $y = h(x)$, такую, что $(x,y) \in \tau(h)$, тривиально: оно должно посимвольно просмотреть входную цепочку x и заменить каждый ее символ a на $h(a)$. Это устройство является примером простейшего транслятора, реализующего трансляцию $\tau(h)$.

Реалистичным примером транслятора, основанного на гомоморфизме, является простейший *ассемблер*.

Транслятором для данной трансляции τ называется такое устройство, которое по данной входной строке x вычисляет выходную цепочку y, такую, что $(x,\ y) \in \tau$. Желательными свойствами транслятора являются

1) *эффективность* (время, затрачиваемое на перевод входной строки, должно быть линейно пропорционально ее длине);

2) малый размер;

3) *правильность* (желательно иметь небольшой тест, такой, чтобы если транслятор прошел его, то это гарантировало бы правильную работу транслятора на *всех* входных цепочках).

§ 1.2. Схемы синтаксически управляемой трансляции

Трансляторы являются средством реализации трансляций, хотя их можно рассматривать также и как способ их задания.

Способом спецификации трансляций, более сложных, чем те, которые описываются при помощи гомоморфизма, является аппарат *схем синтаксически управляемых трансляций* (sdts — syntax-directed translation schema).

***Определение* 1.2.** *Схемой синтаксически управляемой трансляции* называется формальная система $T = (N, \Sigma, \Delta, R, S)$, где N — *алфавит нетерминалов*; Σ — конечный *входной алфавит*; Δ — конечный *выходной алфавит*, причем $N \cap \Sigma = \varnothing$ и $N \cap \Delta = \varnothing$; R — конечное *множество правил схемы* вида $A \to \alpha, \beta$, где $A \in N$, $\alpha \in (N \cup \Sigma)^*$, $\beta \in (N \cup \Delta)^*$, причем каждое вхождение нетерминала в цепочку α взаимно-однозначно связано с некоторым вхождением одноименного нетерминала в β, и эта связь является неотъемлемой частью правила; $S \in N$ — начальный нетерминал.

Пусть $A \to \ \alpha, \beta$ — правило схемы. Цепочка α называется *синтаксической*, а β — *семантической*.

Для указания связей между вхождениями нетерминалов можно использовать индексы. Связанные вхождения нетерминалов помечаются одинаковыми индексами. Например, $A \to aB^{(1)}bCB^{(2)},\ B^{(2)}B^{(1)}dC$.

***Определение* 1.3.** Введем понятие *трансляционной формы* следующим образом:

1) (S, S) — *начальная трансляционная форма*, причем эти два вхождения начального нетерминала связаны друг с другом по определению;

2) если $(\alpha A\beta,\ \alpha' A\beta')$ — трансляционная форма, в которой два явно выделенных вхождения нетерминала A связаны, и если $A \to \gamma, \gamma'$ — правило из R, то $(\alpha\gamma\beta,\ \alpha'\gamma'\beta')$ — трансляционная форма; причем связь между нетерминалами в γ и γ' — такая же, как в правиле; нетерминалы в цепочках α и β связываются с нетерминалами в цепочках α' и β' в новой трансляционной форме точно так же, как в предыдущей; *связь между нетерминалами* является существенной чертой трансляционной формы;

3) трансляционными формами являются такие и только такие пары цепочек, которые могут быть получены с помощью данных двух способов.

Это определение фактически вводит отношение непосредственной выводимости одной трансляционной формы из другой. В таком случае принято писать: $(\alpha A\beta,\ \alpha' A\beta') \underset{T}{\Rightarrow} (\alpha\gamma\beta, \alpha'\gamma'\beta')$. Для степени, транзитивного и рефлексивно-транзитивного замыкания этого отношения используются соответственно следующие обозначения: $\overset{k}{\underset{T}{\Rightarrow}}$, $\overset{+}{\underset{T}{\Rightarrow}}$ и $\overset{*}{\underset{T}{\Rightarrow}}$. Когда ясно, в какой схеме производится вывод, имя схемы может быть опущено.

***Определение* 1.4.** *Трансляция, заданная при помощи схемы синтаксически управляемой трансляции* T, есть множество $\tau(T) = \{(x,\ y) \mid (S, S) \overset{*}{\underset{T}{\Rightarrow}} (x,\ y),\ x \in \Sigma^*,\ y \in \Delta^*\}$ и называется *синтаксически управляемой трансляцией* (sdt).

***Определение* 1.5.** Грамматика $G_{\text{i}} = (N, \Sigma, P_{\text{i}}, S)$, где $P_{\text{i}} = \{A \to \alpha \mid \exists A \to \alpha, \beta \in R\}$, называется *входной грамматикой схемы*. Грамматика $G_{\text{o}} = (N, \Delta, P_{\text{o}}, S)$, где $P_{\text{o}} = \{A \to \beta \mid \exists\ A \to \alpha, \beta \in R\}$, называется *выходной грамматикой схемы*.

Очевидно, что G_i и G_{o} — контекстно-свободные грамматики.

Пример 1.2. Пусть sdts $T = (\{E, T, F\},\ \{a, +, *, (,)\},\ \{a, +, *\},\ R,\ E)$, где

$R = \{(1)\ E \to E{+}T, ET{+};\ (2)\ E \to T, T;\ (3)\ T \to T{*}F, TF{*};$

$(4)\ T \to F, F;\ (5)\ F \to (E), E;\ (6)\ F \to a, a\ \}.$

Связь между нетерминалами в этих правилах очевидна, так как в синтаксической и семантической цепочках каждого правила нет двух или более вхождений одноименных нетерминалов. Рассмотрим какой-нибудь вывод в T, например:

$$(E, E) \Rightarrow (E + T, ET+) \Rightarrow (T^{(1)} + T^{(2)}, T^{(1)}T^{(2)} +) \Rightarrow (F^{(1)} + T^{(2)}, F^{(1)}T^{(2)} +) \Rightarrow$$
$$\Rightarrow (a + T, aT+) \Rightarrow (a + F^{(1)} * F^{(2)}, aF^{(1)}F^{(2)} * +) \Rightarrow$$
$$\Rightarrow (a{+}a{*}F, aaF{*}{+}) \Rightarrow (a{+}a{*}a, aaa{*}{+}).$$

Нетрудно догадаться, что $\tau(T) = \{(x,\ y) \mid x$ — инфиксная запись арифметического выражения, y — эквивалентная постфиксная запись$\}$.

***Определение* 1.6.** Схема синтаксически управляемой трансляции называется *простой*, если в каждом ее правиле $A \to \alpha, \beta$ связанные нетерминалы в цепочках α и β встречаются в одинаковом порядке.

Трансляция, определяемая простой схемой, называется *простой синтаксически управляемой трансляцией*.

Многие, но не все, полезные трансляции могут быть описаны как простые. В примере 1.2 схема T, как и определяемая ею трансляция $\tau(T)$, является простой.

Простые синтаксически управляемые трансляции интересны тем, что каждая из них может быть реализована транслятором в классе недетерминированных магазинных преобразователей (рис. 1.1).

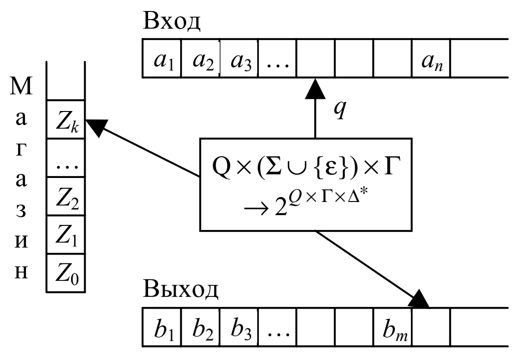

Рис. 1.1.

Другими словами, *магазинные преобразователи характеризуют класс простых синтаксически управляемых трансляций* таким же образом, как магазинные автоматы характеризуют класс контекстно-свободных языков. К рассмотрению таких трансляций мы сейчас и перейдем.

§ 1.3. Магазинные преобразователи и синтаксически управляемые трансляции

В гл. 9 ч. I были описаны обобщенные последовательные машины, называемые также *конечными преобразователями*. Здесь мы рассмотрим магазинные преобразователи, отличающиеся от конечных, лентой памяти, используемой в режиме магазина, как в магазинном автомате.

***Определение* 1.7.** *Магазинный преобразователь* (pdt — pushdown transducer) есть формальная система $P = (Q, \Sigma, \Gamma, \Delta, \delta, q_0, Z_0, F)$, где Q — конечное *множество состояний*, Σ — конечный *входной алфавит*, Γ — конечный *алфавит магазинных символов*, Δ — конечный *выходной алфавит*, $q_0 \in Q$ — *начальное состояние*, $Z_0 \in \Gamma$ — *начальный символ магазина*, $F \subseteq Q$ — множество *конечных состояний*, δ — отображение типа: $Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \to 2^{Q \times \Gamma \times \Delta^*}$.

Запись $\delta(q, a, Z) = \{(p_i, \gamma_i, y_i)\}_{i=1}^{n}$ означает, что pdt P, находясь в состоянии $q \in Q$, сканируя символ $a \in \Sigma \cup \{\varepsilon\}$ на входной ленте и имея $Z \in \Gamma$ на вершине магазина, переходит в одно из состояний $p_i \in Q$, заменяя в магазине символ Z на

$\gamma_i \in \Gamma^*$ и записывая $y_i \in \Delta^*$ на выходную ленту. При этом входная головка сдвигается на одну ячейку вправо, если $a \neq \varepsilon$ (иначе головка остается на месте), головка магазина сканирует последнюю запись в магазине, а головка выходной ленты размещается справа от последней ее записи.

В частности: если $a = \varepsilon$, то выбор действия не зависит от текущего входного символа и, как уже отмечалось, входная головка неподвижна. Если $\gamma_i = \varepsilon$, то верхний символ магазина стирается. Если $y_i = \varepsilon$, то на выходную ленту ничего не пишется, и ее головка остается на прежнем месте.

В начальный момент $q = q_0$, в магазине находится единственный символ Z_0, входная головка сканирует первый входной символ, а выходная лента пуста, причем ее головка находится на первой ячейке.

Работу магазинного преобразователя опишем в терминах конфигураций.

***Определение* 1.8.** *Конфигурацией магазинного преобразователя* P назовем четверку (q, x, α, y), где $q \in Q$ — *текущее состояние*, $x \in \Sigma^*$— часть входной ленты от текущего символа до конца, называемая *непросмотренной частью входной цепочки*, $\alpha \in \Gamma^*$ — *содержимое магазина* (будем считать, что первый символ цепочки α есть *верхний символ магазина*), $y \in \Delta^*$ — *содержимое выходной ленты*. Таким образом, начальная конфигурация есть $(q_0, x, Z_0, \varepsilon)$, где x обозначает всю входную цепочку.

Пусть $(q, ax, Z\alpha, y)$ — текущая конфигурация и $(p, \gamma, z) \in \delta(q, a, Z)$. Тогда один такт работы pdt P записывается как отношение между двумя последовательными конфигурациями: $(q, ax, Z\alpha, y) \vdash (p, x, \gamma\alpha, yz)$. Здесь $q \in Q$, $a \in \Sigma \cup \{\varepsilon\}$, $x \in \Sigma^*$, $Z \in \Gamma$, $\alpha, \gamma \in \Gamma^*$, $y, z \in \Delta^*$.

Как обычно, определяются степень ($\overset{k}{\vdash}$), транзитивное замыкание ($\overset{+}{\vdash}$) и рефлексивно-транзитивное замыкание ($\overset{*}{\vdash}$) этого отношения.

***Определение* 1.9.** Говорят, что $y \in \Delta^*$ есть *выход* для $x \in \Sigma^*$ *при конечном состоянии*, если $(q_0, x, Z_0, \varepsilon) \overset{*}{\vdash} (q, \varepsilon, \alpha, y)$ для некоторых $q \in F$ и $\alpha \in \Gamma^*$.

Трансляция, определяемая магазинным преобразователем P *при конечном состоянии*, есть $\tau(P) = \{(x, y) \mid (q_0, x, Z_0, \varepsilon) \overset{*}{\vdash} (q, \varepsilon, \alpha, y)$ для некоторых $q \in F$ и $\alpha \in \Gamma^*\}$.

Говорят, что $y \in \Delta^*$ есть *выход* для $x \in \Sigma^*$ *при пустом магазине*, если $(q_0, x, Z_0, \varepsilon) \overset{*}{\vdash} (q, \varepsilon, \varepsilon, y)$ для некоторого $q \in Q$.

Трансляция, определяемая магазинным преобразователем P *при пустом магазине*, есть $\tau_e(P) = \{(x, y) \mid (q_0, x, Z_0, \varepsilon) \overset{*}{\vdash} (q, \varepsilon, \varepsilon, y)$ для некоторого $q \in Q\}$.

Пример 10.3. Пусть $P = (\{q\}, \{a, +, *\}, \{E, +, *\}, \{a, +, *\}, \delta, q, E, \{q\})$, где

1) $\delta(q, a, E) = \{(q, \varepsilon, a)\}$,

2) $\delta(q, +, E) = \{(q, EE+, \varepsilon)\}$,

3) $\delta(q, *, E) = \{(q, EE*, \varepsilon)\}$,

4) $\delta(q, \varepsilon, +) = \{(q, \varepsilon, +)\}$,

5) $\delta(q, \varepsilon, *) = \{(q, \varepsilon, *)\}$.

Рассмотрим действия pdt P на входе $+*aaa$:

$$(q, +*aaa, E, \varepsilon) \vdash (q, *aaa, EE+, \varepsilon) \vdash (q, aaa, EE*E+, \varepsilon) \vdash (q, aa, E*E+, a) \vdash$$
$$\vdash (q, a, *E+, aa) \vdash (q, a, E+, aa*) \vdash (q, \varepsilon, +, aa*a) \vdash (q, \varepsilon, \varepsilon, aa*a+).$$

Данный магазинный преобразователь является примером детерминированного магазинного преобразователя. Очевидно, что он преобразует префиксные арифметические выражения в постфиксные.

***Определение* 1.10.** Магазинный преобразователь $P = (Q, \Sigma, \Gamma, \Delta, \delta, q_0, Z_0, F)$ называется *детерминированным* (dpdt) , если

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

2) если $\delta(q, \varepsilon, Z) \neq \varnothing$ для данных $q \in Q$ и $Z \in \Gamma$, то $\delta(q, a, Z) = \varnothing$ для всех $a \in \Sigma$.

На практике предпочитают использовать dpdt, поскольку в реализации они оказываются более эффективными по сравнению с недетерминированными pdt.

Лемма 1.1. *Пусть $T = (N, \Sigma, \Delta, R, S)$ — простая схема синтаксически управляемой трансляции. Существует недетерминированный магазинный преобразователь P, такой, что $\tau_e(P) = \tau(T)$.*

Доказательство. Построим pdt P, о котором идет речь, и покажем, что он реализует трансляцию $\tau(T)$.

Положим $P = (\{q\}, \Sigma, N \cup \Sigma \cup \Delta', \Delta, \delta, q, S, \varnothing)$. Чтобы отличать в магазине P входные символы от выходных, последние переименовываются с помощью гомоморфизма h, определяемого для каждого выходного символа $b \in \Delta$ при помощи равенства $h(b) = b'$ таким образом, чтобы множество символов $\Delta' = \{b' \mid b \in \Delta\}$ не пересекалось со словарем Σ, т.е. $\Sigma \cap \Delta' = \varnothing$.

Отображение δ определяется так:

1. $(q, x_0y_0'B_1x_1y_1' \ldots B_mx_my_m', \varepsilon)$ включается в $\delta(q, \varepsilon, A)$, если $A \to x_0B_1x_1 \ldots B_mx_m, y_0B_1y_1 \ldots B_my_m \in R$, $y_i' = h(y_i)$, $i = 1, 2, \ldots, m$, где $m \ge 0$. Здесь $h(by) = b'h(y)$ для каждого $b \in \Delta$ и $y \in \Delta^*$, $h(\varepsilon) = \varepsilon$.

2. $\delta(q, a, a) = \{(q, \varepsilon, \varepsilon)\}$ для всех $a \in \Sigma$.

3. $\delta(q, \varepsilon, b') = \{(q, \varepsilon, b)\}$ для всех $b \in \Delta$.

I. Докажем сначала, что если $(S, S) \overset{*}{\underset{T}{\Rightarrow}} (x, y)$, то $(q, x, S, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$.

Для этого индукцией по длине вывода l докажем более общее утверждение: если для любого $A \in N$ существует вывод $(A, A) \overset{*}{\underset{T}{\Rightarrow}} (x, y)$, то $(q, x, A, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$.

База. Пусть $l = 1$. Имеем $(A, A) \underset{T}{\Rightarrow} (x, y)$. На этом единственном шаге вывода применяется правило $A \to x, y \in R$. Согласно п.1 определения δ имеем $(q, xy', \varepsilon) \in \delta(q, \varepsilon, A)$. Поэтому $(q, x, A, \varepsilon) \underset{P}{\vdash} (q, x, xy', \varepsilon)$.

Далее согласно п.2 $(q, x, xy', \varepsilon) \overset{|x|}{\underset{P}{\vdash}} (q, \varepsilon, y', \varepsilon)$. Наконец, согласно п.3 имеем $(q, \varepsilon, y', \varepsilon) \overset{|y|}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$. Итак, существует переход $(q, x, A, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$.

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

Индукционный переход. Докажем утверждение для $l = n + 1$.

Пусть $(A, A) \underset{T}{\Rightarrow} (x_0 B_1 x_1 \dots B_m x_m, y_0 B_1 y_1 \dots B_m y_m) \overset{n}{\underset{T}{\Rightarrow}} (x, y)$ — вывод длиной $n + 1$.

Очевидно, что $x = x_0 t_1 x_1 t_2 x_2 \dots t_m x_m$, $y = y_0 z_1 y_1 z_2 y_2 \dots z_m y_m$, (1.1)

причем

$$(B_i, B_i) \overset{l_i}{\underset{T}{\Rightarrow}} (t_i, z_i),\ l_i \le n,\ i = 1, 2, \dots, m. \tag{1.2}$$

На первом шаге данного вывода было применено правило

$$A \to x_0 B_1 x_1 B_2 x_2 \dots B_m x_m,\ y_0 B_1 y_1 B_2 y_2 \dots B_m y_m \in R$$

и потому согласно п.1 построения имеем

$(q, x_0 y_0' B_1 x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', \varepsilon) \in \delta(q, \varepsilon, A)$. (1.3)

Кроме того, согласно индукционной гипотезе из существования частичных выводов (1.2), следует возможность перехода

$(q, t_i, B_i, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, z_i)$, $i = 1, 2, \dots, m$. (1.4)

Рассмотрим движения pdt P. Учитывая условия (1.1) и (1.3), можем написать:

$(q, x, A, \varepsilon) = (q, x_0 t_1 x_1 t_2 x_2 \dots t_m x_m, A, \varepsilon) \underset{P}{\vdash}$

$\underset{P}{\vdash} (q, x_0 t_1 x_1 t_2 x_2 \dots t_m x_m, x_0 y_0' B_1 x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', \varepsilon)$.

Согласно п.2 построений имеем переход

$(q, x_0 t_1 x_1 t_2 x_2 \dots t_m x_m, x_0 y_0' B_1 x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', \varepsilon) \overset{*}{\underset{P}{\vdash}}$

$\overset{*}{\underset{P}{\vdash}} (q, t_1 x_1 t_2 x_2 \dots t_m x_m, y_0' B_1 x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', \varepsilon)$;

согласно п.3 построений имеем переход

$(q, t_1 x_1 t_2 x_2 \dots t_m x_m, y_0' B_1 x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', \varepsilon) \overset{*}{\underset{P}{\vdash}}$

$\overset{*}{\underset{P}{\vdash}} (q, t_1 x_1 \dots t_m x_m, B_1 x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', y_0)$.

Учитывая существование перехода (1.4) для $i = 1$, получаем:

$(q, t_1 x_1 t_2 x_2 \dots t_m x_m, B_1 x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', y_0) \overset{*}{\underset{P}{\vdash}}$

$\overset{*}{\underset{P}{\vdash}} (q, x_1 t_2 x_2 \dots t_m x_m, x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', y_0 z_1)$.

Далее рассуждения с использованием пп.2, 3 построений, а также переходов (1.4) для $i = 2, 3, \dots, m$ повторяются. В результате получаем последующие движения:

$(q, x_1 t_2 x_2 \dots t_m x_m, x_1 y_1' B_2 x_2 y_2' \dots B_m x_m y_m', y_0 z_1) \overset{*}{\underset{P}{\vdash}}$

$\overset{*}{\underset{P}{\vdash}} (q, t_2 x_2 \dots t_m x_m, y_1' B_2 x_2 y_2' \dots B_m x_m y_m', y_0 z_1) \overset{*}{\underset{P}{\vdash}}$

$\overset{*}{\underset{P}{\vdash}} (q, t_2 x_2 \dots t_m x_m, B_2 x_2 y_2' \dots B_m x_m y_m', y_0 z_1 y_1) \overset{*}{\underset{P}{\vdash}} \dots$

$\overset{*}{\underset{P}{\vdash}} (q, t_m x_m, B_m x_m y_m', y_0 z_1 y_1 \dots z_{m-1} y_{m-1}) \overset{*}{\underset{P}{\vdash}}$

$\overset{*}{\underset{P}{\vdash}} (q, x_m, x_m y_m', y_0 z_1 y_1 \dots z_{m-1} y_{m-1} z_m) \overset{*}{\underset{P}{\vdash}}$

$\overset{*}{\underset{P}{\vdash}} (q, \varepsilon, y_m', y_0 z_1 y_1 \dots z_{m-1} y_{m-1} z_m) \overset{*}{\underset{P}{\vdash}}$

$\vdash_P^* (q, \varepsilon, \varepsilon, y_0 z_1 y_1 \dots z_{m-1} y_{m-1} z_m y_m) = (q, \varepsilon, \varepsilon, y)$.

В конце принято во внимание представление цепочки y согласно равенству (1.1).

Итак, вся последовательность движений может быть записана в виде

$$(q, x, A, \varepsilon) \vdash_P^* (q, \varepsilon, \varepsilon, y).$$

В частности, из доказанного вспомогательного утверждения при $A = S$ следует утверждение I.

II. Докажем теперь обратное утверждение:

$$\text{если } (q, x, S, \varepsilon) \vdash_P^* (q, \varepsilon, \varepsilon, y) \text{, то } (S,\ S) \underset{T}{\overset{*}{\Rightarrow}} (x, y).$$

Для этого индукцией по числу l движений типа 1, определенных согласно п.1 построений, докажем более общее утверждение:

если для любого $A \in N$ существует переход

$$(q, x, A, \varepsilon) \vdash_P^* (q, \varepsilon, \varepsilon, y),$$

то

$$(A, A) \underset{T}{\overset{*}{\Rightarrow}} (x, y).$$

База. Пусть $l = 1$. Имеем $(q, x, A, \varepsilon) \vdash_P^* (q, \varepsilon, \varepsilon, y)$, где только одно движение типа 1. Очевидно, что оно — первое движение, так как в исходной конфигурации на вершине магазина находится $A \in N$. Это движение не может привести к появлению нетерминалов в магазинной цепочке из-за того, что они неизбежно привели бы к другим движениям типа 1. Кроме того, магазинная цепочка, замещающая A на вершине магазина, должна начинаться на x, так как только в этом случае удастся продвинуться по входу x (в результате движений, определенных в п.2). Наконец, магазинная цепочка должна заканчиваться на y', потому что только в этом случае на выходе может образоваться цепочка y (в результате движений, определенных в п.3). Поэтому фактически имеем

$$(q, x, A, \varepsilon) \vdash_P (q, x, xy', \varepsilon) \vdash_P^* (q, \varepsilon, y', \varepsilon) \vdash_P^* (q, \varepsilon, \varepsilon, y),$$

где первое движение обусловлено тем, что $(q, xy', \varepsilon) \in \delta(q, \varepsilon, A)$, а это означает существование правила $A \to x, y \in R$. Два последних перехода выполнены согласно пп.2, 3 построений. Воспользовавшись существующим правилом, немедленно получаем вывод $(A, A) \underset{T}{\Rightarrow} (x, y)$.

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

Индукционный переход. Докажем утверждение для $l = n + 1$. Пусть имеется переход $(q, x, A, \varepsilon) \vdash_P^* (q, \varepsilon, \varepsilon, y)$, в котором ровно $n + 1$ движение типа 1. Поскольку в исходной конфигурации на вершине магазина $A \in N$, то первое же движение — типа 1:

$(q, x, A, \varepsilon) \vdash_P (q,\ x,\ x_0 y_0' B_1 x_1 y_1' \dots B_m x_m y_m',\ \varepsilon) \vdash_P^* (q, \varepsilon, \varepsilon, y)$. (1.5)

В конечной конфигурации магазин пуст. Цепочка $x_0 \in \Sigma^*$, появившаяся в верхней части магазина после первого движения, может быть удалена, только если входная цепочка x начинается на x_0. Поэтому далее последуют движения, определяемые п.2, которые продвинут вход по x_0 и удалят такую же цепочку из магазина. Далее ряд движений, определяемых п.3, удалит цепочку y_0' из магазина, выдав на выход y_0, и символ B_1 окажется на вершине магазина. К моменту, когда вершина магазина опустится ниже этой позиции, будет просканирована некоторая часть входа t_1, следующая за цепочкой x_0, а на выходе образуется некоторая цепочка z_1:

$$(q, x, A, \varepsilon) = (q, x_0 t_1 x', A, \varepsilon) \vdash_{P} (q,\ x_0 t_1 x', x_0 y_0' B_1 x_1 y_1' \ldots B_m x_m y_m', \varepsilon) \vdash_{P}^{*}$$
$$\vdash_{P}^{*} (q,\ t_1 x', y_0' B_1 x_1 y_1' \ldots B_m x_m y_m',\ \varepsilon) \vdash_{P}^{*} (q,\ t_1 x', B_1 x_1 y_1' \ldots B_m x_m y_m', y_0) \vdash_{P}^{*}$$
$$\vdash_{P}^{*} (q,\ x', x_1 y_1' \ldots B_m x_m y_m',\ y_0 z_1).$$

Далее мы можем повторить рассуждения, аналогичные рассуждениям предыдущего абзаца, относя их к цепочкам $x_i \in \Sigma^*$, $y_i' \in \Delta'$ $(i = 1, 2, \ldots, m)$ и $B_j \in N$ $(j = 2, \ldots, m)$, последовательно появляющимся в верхней части магазина в результате серии движений, построенных в соответствии с п.2, затем с п.3, и ряда движений, приводящих к понижению вершины магазина ниже позиции, занимаемой B_j. Другими словами, детальный разбор возможных движений от исходной конфигурации к конечной дает основание утверждать, что вход x представим в виде

$$x = x_0 t_1 x_1 \ldots t_m x_m,\ y = x_0 t_1 x_1 \ldots t_m x_m, \tag{1.6}$$

причем

$$(q, t_i, B_i, \varepsilon) \vdash_{P}^{l_i} (q, \varepsilon, \varepsilon, z_i),\ l_i \leq n,\ i = 1, 2, \ldots, m. \tag{1.7}$$

По построению первое движение (1.4) обусловлено существованием правила

$$A \to x_0 B_1 x_1 \ldots B_m x_m, y_0 B_1 y_1 \ldots B_m y_m \in R, \tag{1.8}$$

а из существования движений (1.7) по индукционному предположению следует существование выводов

$$(B_i, B_i) \overset{*}{\underset{T}{\Rightarrow}} (t_i, z_i),\ i = 1, 2, \ldots, m. \tag{1.9}$$

Используя движения (1.8) и (1.9), с учетом (1.6) получаем:

$$(A, A) \underset{T}{\Rightarrow} (x_0 B_1 x_1 \ldots B_m x_m, y_0 B_1 y_1 \ldots B_m y_m) \overset{*}{\underset{T}{\Rightarrow}} (x_0 t_1 x_1 \ldots t_m x_m, y_0 z_1 y_1 \ldots z_m y_m) = (x, y).$$

В частности, при $A = S$ следует утверждение II.

Из рассуждений I и II следует утверждение леммы.

Доказанная лемма дает алгоритм построения недетерминированного магазинного преобразователя, эквивалентного данной простой схеме синтаксически управляемой трансляции.

Пример 1.4. Пусть sdts $T = (\{E\}, \{a, +, *\}, \{a, +, *\}, R, E)$, где

$R = \{(1)\ E \to +EE, EE+\ ;\ (2)\ E \to *EE, EE*\ ;\ (3)\ E \to a, a\}$.

По лемме 10.1 эквивалентный магазинный преобразователь есть

$P = (\{q\}, \{a, +, *\}, \{E, a, +, *, a', +', *'\}, \{a, +, *\}, \delta, q, E, \varnothing)$, где

1) $\delta(q, \varepsilon, E) = \{(q, +EE+', \varepsilon), (q, *EE*', \varepsilon), (q, aa', \varepsilon)\}$,

2) $\delta(q, b, b) = \{(q, \varepsilon, \varepsilon)\}$ для всех $b \in \{a, +, *\}$,

3) $\delta(q, \varepsilon, c') = \{(q, \varepsilon, c)\}$ для всех $c \in \{a, +, *\}$.

Сравните этот недетерминированный магазинный преобразователь с эквивалентным детерминированным pdt из примера 1.3. Оба преобразуют префиксные арифметические выражения в постфиксные.

Лемма 1.2. *Пусть* $P = (Q, \Sigma, \Gamma, \Delta, \delta, q_0, Z_0, \varnothing)$ *— недетерминированный магазинный преобразователь. Существует простая схема синтаксически-управляемой трансляции* T*, такая, что* $\tau(T) = \tau_e(P)$.

Доказательство. Построим такую схему T следующим образом. Положим $T = (N, \Sigma, \Delta, R, S)$, где $N = \{S\} \cup \{[qZp] \mid q, p \in Q, Z \in \Gamma\}$, Σ и Δ — такие же, как в pdt P,

$R = \{S \to [q_0Z_0p], [q_0Z_0p] \mid \text{для всех } p \in Q\} \cup$

$\cup \{[qZp] \to a[q_1Z_1q_2][q_2Z_2q_3]\dots[q_mZ_mq_{m+1}], y[q_1Z_1q_2][q_2Z_2q_3]\dots[q_mZ_mq_{m+1}] \mid$
$(q_1, Z_1Z_2\dots Z_m, y) \in \delta(q, a, Z);\ a \in \Sigma \cup \{\varepsilon\},\ y \in \Delta^*;\ p, q, q_i \in Q; Z, Z_i \in \Gamma;$

$i = 1, 2, \dots, m+1;\ q_{m+1} = p\}$.

I. Докажем сначала, что если $(q, x, Z, \varepsilon) \overset{*}{\underset{P}{\vdash}} (p, \varepsilon, \varepsilon, y)$, то $([qZp], [qZp]) \overset{*}{\underset{T}{\Rightarrow}} (x, y)$, используя индукцию по числу l движений магазинного преобразователя P.

База. Пусть $l = 1$. Имеем $(q, x, Z, \varepsilon) \underset{P}{\vdash} (p, \varepsilon, \varepsilon, y)$. В этом случае $x \in \Sigma \cup \{\varepsilon\}$ и $(p, \varepsilon, y) \in \delta(q, x, Z)$. Тогда по построению схемы T существует правило $[qZp] \to x, y \in R$, с помощью которого немедленно получаем: $([qZp], [qZp]) \underset{T}{\Rightarrow} (x, y)$.

Индукционная гипотеза. Предположим, что утверждение I выполняется для всех переходов между конфигурациями за число движений $l \le n$ ($n \le 1$).

Индукционный переход. Докажем, что тогда утверждение I справедливо и для $l = n + 1$. Итак, пусть $(q, x, Z, \varepsilon) \overset{n+1}{\underset{P}{\vdash}} (p, \varepsilon, \varepsilon, y)$. Рассмотрим этот переход подробнее.

В общем случае первое движение имеет вид

$$(q, x, Z, \varepsilon) = (q, ax', Z, \varepsilon) \underset{P}{\vdash} (q_1, x', Z_1Z_2\dots Z_m, y_0). \tag{1.10}$$

Затем следуют дальнейшие движения:

$$(q_1, x', Z_1Z_2\dots Z_m, y_0) = (q_1, x_1x_2\dots x_m, Z_1\dots Z_m, y_0) \overset{l_1}{\underset{P}{\vdash}} (q_2, x_2\dots x_m, Z_2\dots Z_m, y_0y_1) \overset{l_2}{\underset{P}{\vdash}} \dots$$
$$\dots \overset{l_m}{\underset{P}{\vdash}} (q_{m+1}, \varepsilon, \varepsilon, y_0y_1y_2\dots y_m) = (p, \varepsilon, \varepsilon, y).$$

Здесь $a \in \Sigma \cup \{\varepsilon\}$, $x = ax_1x_2\dots x_m$, $y = y_0y_1y_2\dots y_m$, причем

$$(q_i, x_i, Z_i, \varepsilon) \overset{l_i}{\underset{P}{\vdash}} (q_{i+1}, \varepsilon, \varepsilon, y_i),\ i = 1, 2, \dots, m;\ l_i \le n;\ q_{m+1} = p. \tag{1.11}$$

Первое движение (1.10) существует потому, что $(q_1, Z_1\dots Z_m, y_0) \in \delta(q, a, Z)$, следовательно, по способу построения правил схемы в ней имеется правило

$$[qZp] \to a[q_1Z_1q_2][q_2Z_2q_3]\dots[q_mZ_mq_{m+1}],\ y_0[q_1Z_1q_2][q_2Z_2q_3]\dots[q_mZ_mq_{m+1}], \tag{1.12}$$

в обозначениях нетерминалов которого участвуют те состояния, по которым проходил pdt P. Из последующих движений (1.12) согласно индукционной гипотезе следует существование выводов

$$([q_iZ_iq_{i+1}], [q_iZ_iq_{i+1}]) \overset{*}{\underset{T}{\Rightarrow}} (x_i, y_i),\ i = 1, 2, \dots, m. \tag{1.13}$$

Из (1.12) и (1.13) можно выстроить требуемый вывод:

$$([qZp], [qZp]) \underset{T}{\Rightarrow} (a[q_1Z_1q_2]\dots[q_mZ_mq_{m+1}], y_0[q_1Z_1q_2]\dots[q_mZ_mq_{m+1}]) \overset{*}{\underset{T}{\Rightarrow}}$$

$\overset{*}{\underset{T}{\Rightarrow}} (ax_1x_2\ldots x_m, y_0y_1\ldots y_m) = (x, y)$.

II. Индукцией по длине l вывода докажем теперь обратное утверждение: если $([qZp], [qZp]) \overset{*}{\underset{T}{\Rightarrow}} (x, y)$, то $(q, x, Z, \varepsilon) \overset{*}{\underset{P}{\vdash}} (p, \varepsilon, \varepsilon, y)$.

База. Пусть $l = 1$. Имеем $([qZp], [qZp]) \underset{T}{\Rightarrow} (x, y)$. На единственном шаге этого вывода использовано правило схемы $[qZp] \to x, y$, которое обязано своим происхождением тому, что $(p, \varepsilon, y) \in \delta(q, x, Z)$, где $x \in \Sigma \cup \{\varepsilon\}$, $y \in \Delta^*$, а тогда $(q, x, Z, \varepsilon) \underset{P}{\vdash}$ $(p, \varepsilon, \varepsilon, y)$.

Индукционная гипотеза. Предположим, что аналогичное утверждение выполняется для всех выводов, длина которых не превосходит n $(0 \leq n \leq 1\ )$.

Индукционный переход. Докажем аналогичное утверждение для выводов длиной $l = n + 1$. Пусть

$$([qZp], [qZp]) \underset{T}{\Rightarrow} (a[q_1Z_1q_2]\ldots[q_mZ_mq_{m+1}], y_0[q_1Z_1q_2]\ldots[q_mZ_mq_{m+1}]) \overset{n}{\underset{T}{\Rightarrow}} (x, y). \quad (1.14)$$

Из (1.14) следует, что существует правило

$$[qZp] \to a[q_1Z_1q_2]\ldots[q_mZ_mq_{m+1}],\ y[q_1Z_1q_2]\ldots[q_mZ_mq_{m+1}] \in R,$$

которое обязано своим происхождение тому, что

$$(q_1, Z_1\ldots Z_m,\ y) \in \delta(q, a, Z). \quad (1.15)$$

Кроме того, из (1.14) следует, что

$$x = ax_1\ldots x_m,\ y = y_0y_1\ldots y_m, \quad (1.16)$$

$$([q_iZ_iq_{i+1}], [q_iZ_iq_{i+1}]) \overset{l_i}{\underset{T}{\Rightarrow}} (x_i, y_i),\ l_i \leq n,$$

а тогда согласно индукционной гипотезе

$$(q_i, x_i, Z_i, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q_{i+1}, \varepsilon, \varepsilon, y_i),\ i = 1, 2, \ldots, m;\ q_{m+1} = p. \quad (1.17)$$

Из (1.15) и (1.17) с учетом (1.16) выстраивается последовательность движений:

$$(q, x, Z, \varepsilon) = (\ q,\ ax_1\ldots x_m, Z, \varepsilon) \underset{P}{\vdash} (q_1, x_1\ldots x_m, Z_1\ldots Z_m, y_0) \overset{*}{\underset{P}{\vdash}}$$
$$\overset{*}{\underset{P}{\vdash}} (p, \varepsilon, \varepsilon, y_0y_1\ldots y_m) = (p, \varepsilon, \varepsilon, y).$$

Из рассуждений I и II следует, что для любых $q, p \in Q$ и $Z \in \Gamma$ вывод

$$([qZp], [qZp]) \overset{*}{\underset{T}{\Rightarrow}} (x, y)$$

существует тогда и только тогда, когда

$$(q, x, Z, \varepsilon) \overset{*}{\underset{P}{\vdash}} (p, \varepsilon, \varepsilon, y).$$

В частности, это справедливо для $Z = Z_0$.

В то же время при помощи правила вида $S \to [q_0Z_0p], [q_0Z_0p]$ всегда можно пристроить начало к вышеприведенному выводу, чтобы считать доказанным утверждение

$$(S, S) \underset{T}{\Rightarrow} ([q_0Z_0p], [q_0Z_0p]) \overset{*}{\underset{T}{\Rightarrow}} (x, y)$$

тогда и только тогда, когда

$$(q, x, Z_0, \varepsilon) \overset{*}{\underset{P}{\vdash}} (p, \varepsilon, \varepsilon, y).$$

Лемма доказана.

Из лемм 1.1 и 1.2 следует

Теорема 1.1. *Трансляция* $\tau = \tau(T)$, *где* T *— простая схема синтаксически управляемой трансляции тогда и только тогда, когда существует недетерминированный магазинный преобразователь* P, *такой, что* $\tau_e(P) = \tau$.

Пример 1.5. В предыдущем примере по простой sdts $T = (\{E\}, \{a, +, *\}, \{a, +, *\}, R, E)$, где $R = \{(1)\ E \to +EE,\ EE+\ ;\ (2)\ E \to *EE,\ EE*\ ;\ (3)\ E \to a,\ a\}$, был построен эквивалентный магазинный преобразователь

$P = (\{q\}, \{a, +, *\}, \{E, a, +, *, a', +', *'\}, \{a, +, *\}, \delta, q, E, \varnothing)$,

где

1) $\delta(q, \varepsilon, E) = \{(q, +EE+', \varepsilon), (q, *EE*', \varepsilon), (q, aa', \varepsilon)\}$,

2) $\delta(q, b, b) = \{(q, \varepsilon, \varepsilon)\}$ для всех $b \in \{a, +, *\}$,

3) $\delta(q, \varepsilon, c') = \{(q, \varepsilon, c)\}$ для всех $c \in \{a, +, *\}$.

Теперь по этому недетерминированному преобразователю P мы построим эквивалентную простую схему синтаксически управляемой трансляции, воспользовавшись алгоритмом, описанным в лемме 1.2. Положим

$T = (\{S, [qEq], [q+q], [q*q]\}, \{a, +, *\}, \{a, +, *\}, R, S)$,

$R = \{(1)\ S \to [qEq], [qEq]$;

(2) $[qEq] \to [q+q]\,[qEq]\,[qEq]\,[q+'q],\ [q+q]\,[qEq]\,[qEq]\,[q+'q]$;

(3) $[qEq] \to [q*q]\,[qEq]\,[qEq]\,[q*'q],\ [q*q]\,[qEq]\,[qEq]\,[q*'q]$;

(4) $[qEq] \to [qaq]\,[qa'q],\ [qaq]\,[qa'q]$;

(5) $[qaq] \to a, \varepsilon$;

(6) $[q+q] \to +, \varepsilon$;

(7) $[q*q] \to *, \varepsilon$;

(8) $[qa'q] \to \varepsilon, a$;

(9) $[q+'q] \to \varepsilon, +$;

(10) $[q*'q] \to \varepsilon, *\}$.

Эта схема мало похожа на исходную, в которой было всего три правила. Однако ее можно эквивалентными преобразованиями привести к исходной.

Во-первых, правые части правил $5-10$ можно подставить в правые части правил $2-4$. В результате получим

$R' = \{(1)\ S \to [qEq], [qEq]$;

$(2')\ [qEq] \to +[qEq]\,[qEq],\ [qEq]\,[qEq]+$;

$(3')\ [qEq] \to *[qEq]\,[qEq],\ [qEq]\,[qEq]*$;

$(4')\ [qEq] \to a, a\}$.

Легко видеть, что из (S, S) выводится в точности то же, что и из $([qEq], [qEq])$. Остается заменить в правилах $2'-4'$ слева и справа $[qEq]$ на простое E и отбросить бесполезное правило 1, чтобы получить исходную схему.

Лемма 1.3. *Пусть* $P = (Q, \Sigma, \Gamma, \Delta, \delta, q_0, Z_0, F)$ *— недетерминированный магазинный преобразователь и* $\tau = \tau(P)$. *Существует недетерминированный магазинный преобразователь* P', *такой, что* $\tau_e(P') = \tau$.

Доказательство. Построим pdt P', исходя из того соображения, что он будет моделировать pdt P до тех пор, пока тот не примет свою входную цепочку, а затем pdt P' будет опустошать свой магазин, совершая ε-движения и не выдавая ничего на выход. Таким образом, pdt P' будет принимать то же самое множество входных цепочек при пустом магазине, какое pdt P принимает при конечных состояниях, и выдавать на выход такие же выходные цепочки.

Итак, положим $P' = (Q', \Sigma, \Gamma', \Delta, \delta', q_0', Z_0', \varnothing)$, где входной и выходной алфавиты — такие же, как у pdt P, а множество конечных состояний, которое в этом случае несущественно, пусто; $Q' = Q \cup \{q_0', q_e\}$, $q_0', q_e \notin Q$; $\Gamma' = \Gamma \cup \{Z_0'\}$, $Z_0' \notin \Gamma$.

Отображение δ' определяется следующим образом:

1. $\delta'(q_0', \varepsilon, Z_0') = \{(q_0, Z_0Z_0', \varepsilon)\}$ воспроизводит начальную конфигурацию недетерминированного магазинного преобразователя P.

2. $\delta'(q, a, Z)$ содержит все элементы $\delta(q, a, Z)$ для $q \in Q$, $a \in \Sigma \cup \{\varepsilon\}$, $Z \in \Gamma$, — реализуется собственно моделирование движений pdt P.

3. $\delta'(q, \varepsilon, Z)$ содержит $(q_e, \varepsilon, \varepsilon)$, если $q \in F$, $Z \in \Gamma'$, — происходит переход в состояние-“ловушку”.

4. $\delta'(q_e, \varepsilon, Z) = \{(q_e, \varepsilon, \varepsilon)\}$ для всех $Z \in \Gamma'$ — производится опустошение магазина.

I. Докажем сначала, что если $(x, y) \in \tau(P)$, то $(x, y) \in \tau_e(P')$.

Пусть $(x, y) \in \tau(P)$, т.е. $(q_0, x, Z_0, \varepsilon) \vdash_P^* (q, \varepsilon, \alpha, y)$, где $q \in F$. Посмотрим, как будет действовать pdt P' на таком же входе.

Согласно п.1 построения имеем $(q_0', x, Z_0', \varepsilon) \vdash_{P'} (q_0, x, Z_0Z_0', \varepsilon)$. Далее согласно п.2 pdt P' повторяет все движения pdt P, т.е. $(q_0, x, Z_0Z_0', \varepsilon) \vdash_{P'}^* (q, \varepsilon, \alpha Z_0', y)$, где $q \in F$, а тогда согласно п.2 pdt P' переходит в состояние q_e, оставаясь в котором, в соответствии с п.3 опустошает свой магазин: $(q, \varepsilon, \alpha Z_0', y) \vdash_{P'}^* (q_e, \varepsilon, \varepsilon, y)$. Следовательно, $(x, y) \in \tau(P')$.

II. Докажем теперь обратное, т.е. если $(x, y) \in \tau_e(P')$, то $(x, y) \in \tau(P)$. Пусть $(x, y) \in \tau(P')$, т.е. $(q_0', x, Z_0', \varepsilon) \vdash_{P'}^* (q_e, \varepsilon, \varepsilon, y)$. Рассмотрим подробнее его движения. Первое движение согласно п.1 имеет вид

$$(q_0', x, Z_0', \varepsilon) \vdash_{P'} (q_0, x, Z_0Z_0', \varepsilon).$$

Ясно, что для опустошения магазина pdt P' должен оказаться в состоянии q_e: без этого ему не сбросить символ Z_0', находящийся на “дне” магазина. В свою очередь, согласно п.2 pdt P' достигает состояния q_e только после того, как, повторяя движения pdt P согласно п.3, он оказывается в состоянии $q \in F$. С этого момента он не продвигается по входной цепочке, ничего не пишет на выходную ленту и только стирает магазин. Это значит, что к моменту достижения конечного состояния он уже прочитал всю свою входную цепочку x и записал всю свою выходную цепочку y.

Дальнейшие движения pdt P' имеют вид

$(q_0, x, Z_0Z_0', \varepsilon) \vdash^*_{P'} (q, \varepsilon, \alpha Z_0', y) \vdash^*_{P'} (q_e, \varepsilon, \varepsilon, y)$, где $q \in F$, $\alpha \in \Gamma^*$.

Но на участке $(q_0, x, Z_0Z_0', \varepsilon) \vdash^*_{P'} (q, \varepsilon, \alpha Z_0', y)$ pdt P' лишь повторяет движения недетерминированного магазинного преобразователя P:

$$(q_0, x, Z_0, \varepsilon) \vdash^*_{P'} (q, \varepsilon, \alpha, y),\ q \in F,\ \alpha \in \Gamma^*.$$

Следовательно, $(x, y) \in \tau(P)$.

Из рассуждений I и II следует справедливость леммы.

Лемма 1.4. *Пусть* $P = (Q, \Sigma, \Gamma, \Delta, \delta, q_0, Z_0, F)$ *— недетерминированный магазинный преобразователь и* $\tau = \tau_e(P)$. *Существует недетерминированный магазинный преобразователь* P'*, такой, что* $\tau(P') = \tau$.

Доказательство. Построим pdt P', исходя из того соображения, что pdt P' будет моделировать pdt P до тех пор, пока тот не примет свою входную цепочку, опустошив магазин, а затем pdt P' совершит еще одно — ε-движение, не записывающее ничего на выходную ленту, и перейдет в свое конечное состояние, принимая ту же входную цепочку. Разумеется, надо позаботиться о том, чтобы к этому моменту магазин не был пуст, иначе никакое дальнейшее движение не будет возможно. Таким образом, pdt P' будет принимать то же самое множество входных цепочек при конечном состоянии, выдавая те же выходные цепочки, что и pdt P.

Положим pdt $P' = (Q', \Sigma, \Gamma', \Delta, \delta', q_0', Z_0', F')$. Здесь $Q' = Q \cup \{q_0', q_f\}$, $q_0', q_f \notin Q$; входной и выходной алфавиты — такие же, как у pdt P; $\Gamma' = \Gamma \cup \{Z_0'\}$, $Z_0' \notin \Gamma$; $F' = \{q_f\}$.

Отображение δ' определяется следующим образом:

1. $\delta'(q_0', \varepsilon, Z_0') = \{(q_0, Z_0Z_0', \varepsilon)\}$ воспроизводит начальную конфигурацию недетерминированного магазинного преобразователя P.

2. $\delta'(q, a, Z)$ содержит все элементы $\delta(q, a, Z)$ для $q \in Q$, $a \in \Sigma \cup \{\varepsilon\}$, $Z \in \Gamma$, — происходит собственно моделирование движений pdt P.

3. $\delta'(q, \varepsilon, Z_0') = \{(q_f, \alpha, \varepsilon)\}$, где $\alpha \in \Gamma^*$ — произвольная магазинная цепочка, — определяет переход в конечное состояние.

I. Докажем сначала, что если $(x, y) \in \tau_e(P)$, то $(x, y) \in \tau(P')$. Пусть $(x, y) \in \tau_e(P)$, т.е. $(q_0, x, Z_0, \varepsilon) \vdash^*_P (q, \varepsilon, \varepsilon, y)$. Посмотрим, как будет действовать pdt P' с такой же входной цепочкой.

Соглсно п.1 $(q_0', x, Z_0', \varepsilon) \vdash_{P'} (q_0, x, Z_0Z_0', \varepsilon)$. Далее согласно п.2 pdt P' повторяет все движения pdt P, т.е. $(q_0, x, Z_0Z_0', \varepsilon) \vdash_{P'}^{*} (q, \varepsilon, Z_0', y)$, а затем согласно п.3 pdt P' переходит в состояние q_{f}: $(q, \varepsilon, Z_0', y) \vdash_{P'}^{*} (q_{\mathrm{f}}, \varepsilon, \alpha, y)$. Следовательно, $(x, y) \in \tau(P')$.

II. Докажем теперь обратное, т.е. если $(x,y) \in \tau(P')$, то $(x, y) \in \tau_{\mathrm{e}}(P)$. Пусть $(x, y) \in \tau(P')$, т.е. $(q_0', x, Z_0', \varepsilon) \vdash_{P'}^{*} (q_{\mathrm{f}}, \varepsilon, \alpha, y)$ при некотором $\alpha \in \Gamma^*$. Рассмотрим подробнее его движения. Первое движение согласно п.1 имеет вид

$$(q_0', x, Z_0', \varepsilon) \vdash_{P'} (q_0, x, Z_0Z_0', \varepsilon).$$

Далее pdt P' оказывается в своем конечном состоянии q_{f}. Это возможно только в результате движения, построенного в соответствии с п.3, который применим лишь тогда, когда на вершине магазина pdt P' покажется символ Z_0', т.е. дальнейшие движения имеют вид

$$(q_0, x, Z_0Z_0', \varepsilon) \vdash_{P'}^{*} (q, \varepsilon, Z_0', y) \vdash_{P'} (q_{\mathrm{f}}, \varepsilon, \alpha, y).$$

На участке $(q_0, x, Z_0Z_0', \varepsilon) \vdash_{P'}^{*} (q, \varepsilon, Z_0', y)$ магазинный преобразователь P' просто повторяет движения pdt P: $(q_0, x, Z_0, \varepsilon) \vdash_{P}^{*} (q, \varepsilon, \varepsilon, y)$. А это значит, что $(x, y) \in \tau_{\mathrm{e}}(P)$.

Из рассуждений I и II следует утверждение леммы.

Из лемм 1.3 и 1.4 следует

Теорема 1.2. *Классы трансляций, реализуемых недетерминированными магазинными преобразователями при конечном состоянии и пустом магазине, совпадают.*

Замечание 1.1. Принимая во внимание теорему 1.2, формулировку теоремы 1.1 можно усилить, не подчеркивая того факта, что простая трансляция реализуется недетерминированным магазинным преобразователем при пустом магазине.

К сожалению, не всякая, даже простая, синтаксически управляемая трансляция может быть реализована *детерминированным* МП-преобразователем.

§ 1.4. Детерминированная генерация выходной цепочки трансляции по левостороннему анализу входной цепочки

***Определение* 1.11.** Схема синтаксически управляемой трансляции называется *семантически однозначной*, если в ней не существует двух правил вида $A \to \alpha, \beta$ и $A \to \alpha, \gamma$, в которых $\beta \neq \gamma$.

Другими словами, семантическая цепочка однозначно определяется правилом входной грамматики.

***Определение* 1.12.** Пусть $G = (V_{\mathrm{N}}, V_{\mathrm{T}}, P, S)$ — контекстно-свободная грамматика, правила которой занумерованы как $1, 2, \ldots, p$ и $S \overset{\pi}{\underset{\mathrm{lm}}{\Rightarrow}} x$ — левосторонний вывод $x \in V_{\mathrm{T}}^*$ в грамматике G. Последовательность номеров правил π, примененных в этом выводе, называется *левосторонним анализом цепочки x*.

Теорема 1.3. *Пусть* $T = (N, \Sigma, \Delta, R, S)$ *— простая семантически однозначная схема синтаксически управляемой трансляции, правила которой занумерованы как* $1, 2, \ldots, p$. *Существует детерминированный магазинный преобразователь* P, *такой, что* $\tau_e(P) = \{(\pi, y) \mid (S, S) \underset{\mathrm{lm}}{\overset{\pi}{\Rightarrow}} (x, y)$ *для некоторой цепочки* $x \in \Sigma^*\}$.

Говоря неформально, выходная цепочка трансляции y может быть сгенерирована детерминированным магазинным преобразователем (dpdt) по левостороннему анализу π входной цепочки x.

Доказательство. Построим dpdt P, о котором идет речь, положив

$$P = (\{q\}, \{1, 2, \ldots, p\}, N \cup \Delta, \Delta, \delta, q, S, \varnothing),$$

где δ определяется следующим образом:

1. $\delta(q, i, A) = (q, \beta, \varepsilon)$, если $A \to \alpha, \beta$ — i-е правило схемы, единственное, начинающееся на $A \to \alpha$.

2. $\delta(q, \varepsilon, b) = (q, \varepsilon, b)$ для всех $b \in \Delta$.

Магазинный преобразователь P — детерминирован, так как правила вида 1 применимы, только если на вершине магазина находится нетерминал, а правила вида 2 используются только тогда, когда на вершине магазина находится выходной символ. Остается доказать, что построенный dpdt P действительно реализует требуемую трансляцию.

I. Индукцией по длине вывода l покажем, что если для любого $A \in N$ существует левосторонний вывод $(A, A) \underset{\mathrm{lm}}{\overset{\pi}{\Rightarrow}} (x, y)$, то $(q, \pi, A, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$.

База. Пусть $l = 1$. На единственном шаге вывода $(A, A) \underset{\mathrm{lm}}{\overset{i}{\Rightarrow}} (x, y)$ применяется правило схемы $A \to x, y$ с номером i. В соответствии с п.1 построения $\delta(q, i, A) = (q, y, \varepsilon)$, и тогда $(q, i, A, \varepsilon) \underset{P}{\vdash} (q, \varepsilon, y, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$. Последний переход совершен согласно п.2 построений, так как $y \in \Delta^*$.

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

Индукционный переход. Покажем, что утверждение выполняется и для $l = n + 1$.

Пусть $(A, A) \underset{\mathrm{lm}}{\overset{i}{\Rightarrow}} (x_0 A_1 x_1 A_2 \ldots A_m x_m,\ y_0 A_1 y_1 A_2 \ldots A_m y_m) \underset{\mathrm{lm}}{\overset{\pi'}{\Rightarrow}} (x, y)$ — вывод длиной $n + 1$; $|\pi'| = n$; $A, A_j \in N$; $x_j \in \Sigma^*$, $y_j \in \Delta^*$, $j = 1, 2, \ldots, m$. На первом шаге применяется правило схемы $A \to x_0 A_1 x_1 A_2 \ldots A_m x_m, y_0 A_1 y_1 A_2 \ldots A_m y_m$, имеющее номер i. Согласно п.1

$$\delta(q, i, A) = (q, y_0 A_1 y_1 A_2 \ldots A_m y_m, \varepsilon). \tag{1.18}$$

Одновременно ясно, что

$$x = x_0 t_1 x_1 t_2 \ldots t_m x_m, \quad y = y_0 z_1 y_1 z_2 \ldots z_m y_m, \quad \pi = i \pi_1 \pi_2 \ldots \pi_m, \tag{1.19}$$

где $(A_j, A_j) \underset{\mathrm{lm}}{\overset{\pi_j}{\Rightarrow}} (t_j, z_j)$, $|\pi_j| \le n$, и, следовательно, по индукционной гипотезе

$$(q, \pi_j, A_j, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, z_j) \text{ для } j = 1, 2, \ldots, m. \tag{1.20}$$

Принимая во внимание (1.18)–(1.20), а также п.2, можем написать:

$$(q, \pi, A, \varepsilon) = (q, i\pi_1\pi_2 \ldots \pi_m, A, \varepsilon) \underset{P}{\vdash} (q, \pi_1\pi_2 \ldots \pi_m, y_0 A_1 y_1 A_2 \ldots A_m y_m, \varepsilon) \overset{*}{\underset{P}{\vdash}}$$
$$\overset{*}{\underset{P}{\vdash}} (q, \pi_1\pi_2 \ldots \pi_m, A_1 y_1 A_2 \ldots A_m y_m, y_0) \overset{*}{\underset{P}{\vdash}} (q, \pi_2 \ldots \pi_m, y_1 A_2 \ldots A_m y_m, y_0 z_1) \overset{*}{\underset{P}{\vdash}} \ldots$$
$$\ldots \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, y_m, y_0 z_1 y_1 \ldots z_m) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y_0 z_1 y_1 \ldots z_m y_m) = (q, \varepsilon, \varepsilon, y).$$

II. Индукцией по длине ввода $l = |\pi|$ покажем, что для любого $A \in N$, если $(q, \pi, A, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$, то $(A, A) \underset{\mathrm{lm}}{\overset{\pi}{\Rightarrow}} (x, y)$ при некотором $x \in \Sigma^*$.

База. Пусть $l = 1$, т.е. $\pi = i$, и $(q, i, A, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$. В общем случае этот переход может происходить только следующим образом: $(q, i, A, \varepsilon) \underset{P}{\vdash} (q, \varepsilon, y, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$. Действительно, первое движение происходит согласно п.1, поскольку на вершине магазина $A \in N$. Других движений этого типа не существует, так как каждое из них продвигается по входной цепочке на один символ. Возможные же движения согласно п.2 предполагают нахождение в верхней части магазина некоторой цепочки $y \in \Delta^*$.

Первое движение определяется значением $\delta(q, i, A) = (q, y, \varepsilon)$, предполагающим существование правила вида $A \to x, y \in R$ с номером i при некотором $x \in \Sigma^*$. С его помощью немедленно получаем $(A, A) \underset{\mathrm{lm}}{\overset{i}{\Rightarrow}} (x, y)$.

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

Индукционный переход. Предположим, что утверждение выполняется и для $l = n + 1$. Пусть $(q, \pi, A, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y)$, где $|\pi| = n + 1$. В общем случае эти движения могут происходить только следующим образом:

$$(q, \pi, A, \varepsilon) = (q, i\pi', A, \varepsilon) \underset{P}{\vdash} (q, \pi', y_0A_1y_1A_2\ldots A_my_m, \varepsilon) \overset{*}{\underset{P}{\vdash}}$$
$$\overset{*}{\underset{P}{\vdash}} (q, \pi', A_1y_1A_2\ldots A_my_m, y_0) = (q, \pi_1\pi_2\ldots\pi_m, A_1y_1A_2\ldots A_my_m, y_0) \overset{*}{\underset{P}{\vdash}}$$
$$\overset{*}{\underset{P}{\vdash}} (q, \pi_2\ldots\pi_m, y_1A_2\ldots A_my_m, y_0z_1) \overset{*}{\underset{P}{\vdash}} (q, \pi_2\ldots\pi_m, A_2\ldots A_my_m, y_0z_1y_1) \overset{*}{\underset{P}{\vdash}} \ldots$$
$$\ldots \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, y_m, y_0z_1y_1\ldots z_m) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, y_0z_1y_1\ldots z_my_m) = (q, \varepsilon, \varepsilon, y).$$

Соответственно

$$\pi = i\pi_1\pi_2\ldots\pi_m,\ \ y = y_0z_1y_1z_2\ldots z_my_m. \tag{1.21}$$

Первое движение позволяет утверждать, что $\delta(q, i, A) = (q, y_0A_1y_1A_2\ldots A_my_m, \varepsilon)$, и, следовательно, в схеме существует правило под номером i вида

$$A \to x_0A_1x_1A_2\ldots A_mx_m, y_0A_1y_1A_2\ldots A_my_m. \tag{1.22}$$

Промежуточные движения вида $(q, \pi_j, A_j, \varepsilon) \overset{*}{\underset{P}{\vdash}} (q, \varepsilon, \varepsilon, z_j)$, $|\pi_j| \le n$, по индукционному предположению гарантируют существование выводов вида

$$(A_j, A_j) \underset{\mathrm{lm}}{\overset{\pi_j}{\Rightarrow}} (t_j, z_j) \text{ при некоторых } t_j \in \Delta^*,\ \ j = 1, 2, \ldots, m. \tag{1.23}$$

Используя правило (1.22), выводы (1.23) и учитывая равенство (1.21), можем построить вывод

$$(A, A) \underset{\mathrm{lm}}{\overset{i}{\Rightarrow}} (x_0A_1x_1A_2\ldots A_mx_m, y_0A_1y_1A_2\ldots A_my_m) \underset{\mathrm{lm}}{\overset{\pi'}{\Rightarrow}}$$
$$\underset{\mathrm{lm}}{\overset{\pi'}{\Rightarrow}} (x_0t_1x_1t_2\ldots t_mx_m, y_0z_1y_1z_2\ldots z_my_m) = (x, y).$$

Здесь $\pi' = \pi_1\pi_2\ldots\pi_m$, $i\pi' = \pi$, т.е. $(A, A) \underset{\mathrm{lm}}{\overset{\pi}{\Rightarrow}} (x, y)$.

Утверждение теоремы следует из рассуждений I и II при $A = S$.

Пример 1.6. Пусть $T = (\{E, T, F\}, \{a, +, *, (,)\}, \{a, +, *\}, R, E)$, где

$R = \{$(1) $E \to E + T, +ET$;

(2) $E \to T, T$;

(3) $T \to T * F, *TF$;

(4) $T \to F, F$;

(5) $F \to (E), E$;

(6) $F \to a, a\ \}$.

В соответствии с описанием построений в теореме 1.3 определим детерминированный магазинный преобразователь, восстанавливающий выход трансляции по левостороннему анализу входной цепочки:

$$P = \big(\{q\}, \{1, 2, 3, 4, 5, 6\}, (\{E, T, F, a, +, *\}, \{a, +, *\}, \delta, q, E, \varnothing\big),$$

где

1) $\delta(q, 1, E) = (q, +ET, \varepsilon)$;
2) $\delta(q, 2, E) = (q, T, \varepsilon)$;
3) $\delta(q, 3, T) = (q, *TF, \varepsilon)$;
4) $\delta(q, 4, T) = (q, F, \varepsilon)$;
5) $\delta(q, 5, F) = (q, E, \varepsilon)$;
6) $\delta(q, 6, F) = (q, a, \varepsilon)$;
7) $\delta(q, \varepsilon, a) = (q, \varepsilon, a)$;
8) $\delta(q, \varepsilon, +) = (q, \varepsilon, +)$;
9) $\delta(q, \varepsilon, *) = (q, \varepsilon, *)$.

Очевидно, что при $\pi = 23465124646$ имеем $(E, E) \underset{lm}{\overset{\pi}{\Rightarrow}} (a * (a + a), * a + aa)$. Нетрудно убедиться в том, что $(q, 23465124646, E, \varepsilon) \mathrel{\vdash_P^*} (q, \varepsilon, \varepsilon, * a + aa)$.