

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
УФИМСКИЙ ГОСУДАРСТВЕННЫЙ АВИАЦИОННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Кафедра ТК

курс лекций по дисциплине

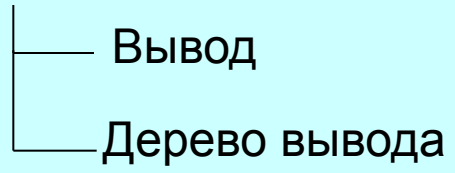
Методы построения трансляторов

Тема: Цепочки вывода

Преподаватель: к.т.н., доцент Карамзина А.Г.

Тема № 5

Цепочки вывода



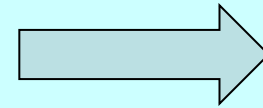
Вывод

процесс порождения предложения языка на основе правил определяющей язык грамматики

Цепочка $\beta = \delta_1 \gamma \delta_2$ *непосредственно выводима* из цепочки $\alpha = \delta_1 \omega \delta_2$

в грамматике $G(VT, VN, P, S)$, $V = VT \cup VN$, $\delta_1, \gamma, \delta_2 \in V^*$, $\omega \in V^+$,

если в грамматике G существует правило: $\omega \rightarrow \gamma \in P$



$\alpha \Rightarrow$
 β

Цепочка β *выводима* из цепочки α (обозначается: $\alpha \Rightarrow^* \beta$), если выполняется одно из двух условий:

- β непосредственно выводима из α ($\alpha \Rightarrow \beta$);
- $\exists \gamma$ такая, что: γ выводима из α и β непосредственно выводима из γ ($\alpha \Rightarrow^* \gamma$ и $\gamma \Rightarrow \beta$).

Вывод

Последовательность непосредственно выводимых цепочек называется выводом или *цепочкой вывода*.

Каждый шаг
в цепочке

шагов вывода в цепочке вывода всегда на один больше, чем промежуточных цепочек – если цепочка β непосредственно выводима из цепочки α : $\alpha \Rightarrow \beta$, то имеется всего один шаг вывода

Если цепочка β выводима из цепочки α за n шагов вывода, то она обозначается так:

$$\alpha \Rightarrow^+ \beta \text{ или } \alpha \Rightarrow^* \beta$$

цепочка β нетривиально выводима из цепочки α .

Если количество шагов вывода известно, то его можно указать непосредственно,

например, $\alpha \Rightarrow^3 \beta$ – цепочка β выводится из цепочки α за 3 шага вывода.

Вывод

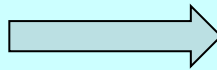
Пример: задана грамматика **G** для целых десятичных чисел со знаком:

$G(\{0,1,2,3,4,5,6,7,8,9,-,+ \}, \{S, T, F\}, P, S)$:

$P: S \rightarrow T \mid +T \mid -T$

$T \rightarrow F \mid TF$

$F \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9.$



$P: S \rightarrow T^1 \mid +T^2 \mid -T^3$

$T \rightarrow F^4 \mid TF^5$

$F \rightarrow$

$0^6 \mid 1^7 \mid 2^8 \mid 3^9 \mid 4^{10} \mid 5^{11} \mid 6^{12} \mid 7^{13} \mid 8^{14} \mid 9^{15}$

$S \Rightarrow^* 575$

$S \xRightarrow{1} T \xRightarrow{5} TF \xRightarrow{11} T5 \xRightarrow{5} TF5 \xRightarrow{13} T75 \xRightarrow{4} F75 \xRightarrow{11} 575$

$S \Rightarrow^7 575$

$-T \Rightarrow^* -9856$

$-T \xRightarrow{5} -TF \xRightarrow{5} -TFF \xRightarrow{5} -TFFF \xRightarrow{4} -FFFF \xRightarrow{15} -9FFF \xRightarrow{11} -9F5F \xRightarrow{14} -985F \xRightarrow{12} -9856$ $-T \Rightarrow^8 -9856$

$S \Rightarrow^* +32$

$S \xRightarrow{2} +T \xRightarrow{5} +TF \xRightarrow{4} +FF \xRightarrow{9} +3F \xRightarrow{8} +32$

$S \Rightarrow^5 +32$

$+FT \Rightarrow^* +104$

$+FT \xRightarrow{5} +FTF \xRightarrow{4} +FFF \xRightarrow{7} +1FF \xRightarrow{6} +10F \xRightarrow{10} +104$

$+FT \Rightarrow^5 +104$

Вывод

Вывод называется **законченным**, если на основе цепочки β , полученной в результате вывода, нельзя больше сделать ни одного шага вывода –

вывод называется законченным, если цепочка β , полученная в результате вывода, *пустая* или содержит только *терминальные символы* грамматики $G(VT, VN, P, S)$: $\beta \in VT^*$.

Цепочка β , полученная в результате законченного вывода, называется **конечной** цепочкой вывода.

Цепочка символов $\alpha \in V^*$ называется **сентенциальной формой грамматики** $G(VT, VN, P, S)$, $V = VT \cup VN$, если она выводима из целевого символа грамматики S :

$$S \Rightarrow^* \alpha.$$

Если цепочка $\alpha \in VT^*$ получена в результате законченного вывода, то она называется **конечной сентенциальной формой**.

Вывод

Вывод называется **левосторонним**, если в нем на каждом шаге вывода правило грамматики применяется всегда к крайнему левому нетерминальному символу в цепочке (на каждом шаге вывода происходит подстановка цепочки символов на основании правила грамматики вместо крайнего левого нетерминального символа в исходной цепочке).

$$S \xrightarrow{2} +T \xrightarrow{5} +TF \xrightarrow{4} +FF \xrightarrow{9} +3F \xrightarrow{8} +32$$

Вывод называется **правосторонним**, если в нем на каждом шаге вывода правило грамматики применяется всегда к крайнему правому нетерминальному символу в цепочке (на каждом шаге вывода происходит подстановка цепочки символов на основании правила грамматики вместо крайнего

$$S \xrightarrow{1} T$$

Для КС-грамматик и регулярных грамматик для любой сентенциальной формы всегда можно построить левосторонний или правосторонний выводы.

Для грамматик других типов это не всегда возможно, так как по структуре их правил не всегда можно выполнить замену крайнего левого и крайнего правого нетерминального символа в цепочке.

Дерево вывода

Дерево вывода грамматики $G(VT, VN, P, S)$ – это дерево (*граф*), которое соответствует некоторой цепочке вывода и удовлетворяет следующим условиям:

- каждая вершина дерева обозначается символом грамматики $A \in (VT \cup VN)$;

По структуре правил дерево вывода в указанном виде всегда можно построить только для КС-грамматик и регулярных грамматик. Для грамматик других типов дерево вывода в таком виде можно построить не всегда (либо же оно будет иметь несколько иной вид).

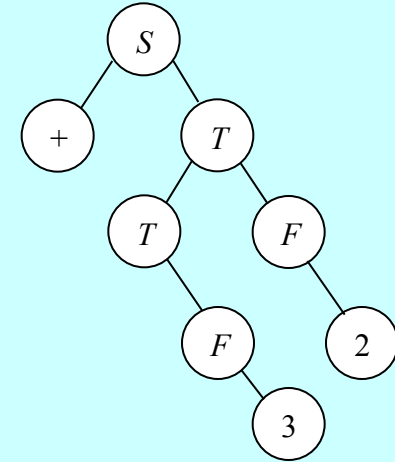
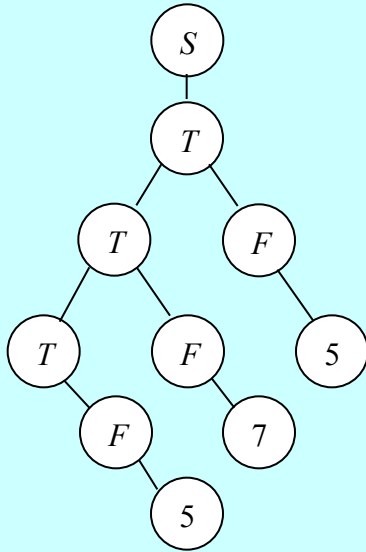
- если некоторый узел дерева обозначен символом $A \in VN$, а облегающие его узлы – символами b_1, b_2, \dots, b_n ; $n > 0$, $\forall n \geq i > 0: b_i \in (VT \cup VN \cup \{\lambda\})$, то в грамматике $G(VT, VN, P, S)$ существует правило $A \rightarrow b_1, b_2, \dots, b_n \in P$.

Дерево вывода

Деревья вывода для цепочек вывода 1 и 3:

$$S \xrightarrow{1} T \xrightarrow{5} TF \xrightarrow{11} T5 \xrightarrow{5} TF5 \xrightarrow{13} T75 \xrightarrow{4} F75 \xrightarrow{11} 575$$

$$S \xrightarrow{2} +T \xrightarrow{5} +TF \xrightarrow{4} +FF \xrightarrow{9} +3F \xrightarrow{8} +32$$



Для того чтобы построить дерево вывода, достаточно иметь цепочку вывода.

Дерево вывода

Дерево вывода можно построить двумя способами:

- *сверху вниз* – построение начинается с целевого символа грамматики, который помещается в корень.

Затем в грамматике выбирается правило, соответствующее выводу корневой символики на первом уровне.

для строго формализованного построения дерева вывода всегда удобнее пользоваться строго определенным выводом: либо левосторонним, либо правосторонним

На втором шаге среди правил грамматики выбирается правило, соответствующее выводу корневой символики на первом уровне. Если выбранное правило левостороннее, то левая вершина (крайняя левая – для левостороннего вывода, крайняя правая – для правостороннего) вершина, обозначенная нетерминальным символом, для этой вершины выбирается нужное правило грамматики, и она раскрывается на несколько вершин следующего уровня.

Построение дерева заканчивается, когда все концевые вершины обозначены терминальными символами, в противном случае надо вернуться ко второму шагу и продолжить построение;

Дерево вывода

- *снизу вверх* – построение начинается с листьев дерева.

В качестве листьев выбираются терминальные символы конечной цепочки вывода, которые на первом шаге построения образуют последний уровень (*слой*) дерева (*построение дерева идет по слоям*).

На втором шаге построения выбирается символ, который является левым сыном выбранного символа. Выбранная операция выбирается вместо операции, которая была выбрана на первом шаге.

Так как все известные языки программирования имеют нотацию записи «слева – направо», компилятор также всегда читает входную программу слева направо (и сверху вниз, если программа разбита на несколько строк), то для построения дерева вывода методом «сверху вниз», как правило, используется левосторонний вывод, а для построения «снизу вверх» – правосторонний вывод.

Построение дерева вывода продолжается до тех пор, пока не будет достигнута корневая целевым символом), а иначе надо вернуться ко второму шагу и повторить процесс выбора операции над полученным слоем дерева.

Дерево вывода

Грамматика называется **однозначной**, если для каждой цепочки символов языка, заданного этой грамматикой, можно построить единственный левосторонний (и единственный правосторонний) вывод или для каждой цепочки символов языка, заданного этой грамматикой, существует единственное дерево вывода (в противном случае грамматика называется **неоднозначной**).

В общем виде невозможно проверить, является ли заданная грамматика однозначной или нет.

Однозначность – это свойство грамматики, а не языка.

Однако для КС-грамматик существуют определенного вида правила, по наличию которых во множестве правил грамматики $G(VT, VN, P, S)$ можно утверждать, что она является неоднозначной (отсутствие правил указанного вида (всех вариантов) есть необходимое, но не достаточное условие однозначности грамматики) $A \in VN; \alpha, \beta, \gamma \in (VN \cup VT)^*$:

$A \rightarrow AA \mid \alpha;$

$A \rightarrow A\alpha A \mid \beta;$

$A \rightarrow \alpha A \mid A\beta \mid \gamma;$

$A \rightarrow \alpha A \mid \alpha A\beta A \mid \gamma.$

Контрольная работа № 2

Грамматика для целых десятичных чисел со знаком задана:

$G(\{0,1,2,3,4,5,6,7,8,9,-,+ \},\{S,T,F\},P,S)$:

P:

$S \rightarrow T \mid +T \mid -T$

$T \rightarrow F \mid TF$

$F \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

Построить дерево вывода для цепочки вывода

<p>Вариант 1</p> <p>$S \Rightarrow 5987$</p> <p>$-T \Rightarrow -555$</p>	<p>Вариант 2</p> <p>$S \Rightarrow 98521$</p> <p>$T \Rightarrow 145$</p>	<p>Вариант 3</p> <p>$S \Rightarrow 65$</p> <p>$+T \Rightarrow +63957$</p>
<p>Вариант 4</p> <p>$S \Rightarrow -965$</p> <p>$-T \Rightarrow -694$</p>	<p>Вариант 5</p> <p>$S \Rightarrow +698$</p> <p>$T \Rightarrow 6974$</p>	<p>Вариант 6</p> <p>$S \Rightarrow 1112$</p> <p>$-T \Rightarrow -787$</p>
<p>Вариант 7</p> <p>$S \Rightarrow -768$</p> <p>$+T \Rightarrow +787$</p>	<p>Вариант 8</p> <p>$S \Rightarrow +621$</p> <p>$-T \Rightarrow -1111$</p>	<p>Вариант 9</p> <p>$S \Rightarrow 6412$</p> <p>$-T \Rightarrow -114$</p>