

Глава II. Семантический анализ. Атрибутная грамматика

Вступление:

- 1) Цепочки = тексты компьютерных программ (в широком смысле).
- 2) Первая цель семантического анализа программ – проверка корректности программы на более глубоком уровне, чем обычный анализ средствами КС-грамматики.

Вторая цель – сбор и систематизация дополнительной («недостающей») информации о программе.
- 3) Атрибутная грамматика – расширение КС-грамматики.

§ 1. Атрибутная грамматика: основные понятия

Напоминание: грамматический символ $\in \Sigma \cup \Gamma$.

Пусть G – фиксированная КС-грамматика.

Опр. Атрибутом грамматического символа X называется параметр, приписанный X , имеющий название (переменная).

Доменом атрибута называется область значений, которые может иметь соответствующий параметр.

Для X может существовать конечное множество атрибутов (может быть пустое).

Для терминальных символов в качестве атрибута используется только его лексическое значение, хранимое в таблице символов. Его обозначение, например, `t.lexval` .

Для нетерминального символа, например `A`, обозначение атрибута `val` записывается `A.val` .

Опр. Семантическим правилом (действием), поставленным в соответствие правилу вывода $A \rightarrow \alpha$, называется функциональная зависимость $b \underbrace{:=}_{\text{присваивание}} f(c_1, \dots, c_k)$, где b, c_1, \dots, c_k – атрибуты символов, принадлежащих $\{A, \alpha\}$.

Для каждого правила вывода может существовать конечное множество семантических правил (может быть пустое).

Выделяют два класса атрибутов: синтезируемые и наследуемые. Это зависит от семантического правила вычисления атрибута.

Опр. Синтезируемый атрибут $A.s$ символа A вычисляется по формуле $A.s := f(c_1, \dots, c_k)$, где действие соответствует правилу вывода $A \rightarrow \alpha$, и c_1, \dots, c_k – атрибуты символов, принадлежащих α .

Опр. Наследуемый атрибут $A.i$ символа A вычисляется по формуле $A.i := f(c_1, \dots, c_k)$, где действие соответствует правилу вывода $B \rightarrow \alpha A \beta$, и c_1, \dots, c_k – атрибуты символов, принадлежащих $\{B, \alpha, \beta\}$.

Опр. Атрибутивной грамматикой называется набор из КС-грамматики G , множества атрибутов и множества семантических правил.

Замечание: фиксированную грамматику G можно преобразовать в атрибутивную многими способами.

Опр. Аннотированным деревом вывода (в атрибутивной грамматике) называется дерево вывода цепочки w , в котором каждому узлу X соответствует список значений атрибутов грамматического символа X .

Замечание:

В аннотированном дереве синтезируемый атрибут $A.s$ зависит от значений атрибутов сыновей узла A . Следовательно, значения всех синтезируемых атрибутов вычисляются «снизу-вверх».

Наследуемый атрибут $A.i$ зависит от значений атрибутов отца узла A (обозначен B), или от значений атрибутов братьев узла A (принадлежат α, β). Следовательно, его значение вычисляется либо «сверху-вниз», либо комбинированным способом обхода дерева.

Пример. (стр. 186 книжки Шура и Замятина)

$G = \{E \rightarrow E+T \mid T, T \rightarrow T * F \mid F, F \rightarrow (E) \mid x\}$

Пусть $E.val$, $T.val$, $F.val$ – значение арифметического выражения;
 $x.lexval$ – лексическое значение x .

Атрибутная грамматика записывается в виде таблицы, в которой правилам вывода соответствуют семантические правила:

№	Правило вывода	Семантические правила

№	Правило вывода	Семантические правила
1	$E \rightarrow E+T$	
2	$E \rightarrow T$	$E.val := T.val$
3	$T \rightarrow T * F$	
4	$T \rightarrow F$	$T.val := F.val$
5	$F \rightarrow (E)$	$F.val := E.val$
6	$F \rightarrow x$	$F.val := x.lexval$

№	Правило вывода	Семантические правила
1	$E \rightarrow E_1 + T$	$E.val := E_1.val + T.val$
2	$E \rightarrow T$	$E.val := T.val$
3	$T \rightarrow T_1 * F$	$T.val := T_1.val \cdot F.val$
4	$T \rightarrow F$	$T.val := F.val$
5	$F \rightarrow (E)$	$F.val := E.val$
6	$F \rightarrow x$	$F.val := x.lexval$

Вопрос: атрибут .val синтезируемый или наследуемый?

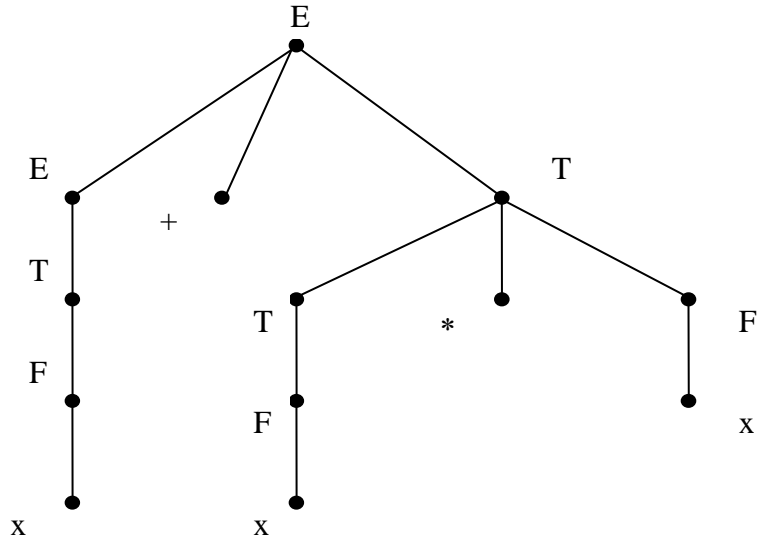
Построим аннотированное дерево вывода арифметического выражения

$$2 + 5 * 3.$$

Ему соответствует цепочка $x + x * x$.

Дерево вывода:

Дерево вывода:



Аннотированное дерево вывода для $2 + 5 * 3$:

