

编译技术



胡春明
hucm@buaa.edu.cn

2019.9-2019.12



编译过程是指将**高级语言程序**翻译为等价的**目标程序**的过程。

习惯上是将编译过程划分为5个基本阶段：



前两节课（符号表+运行时存储）：为代码生成做“需求分析”

符号表
(编译时数据结构)

符号表

	name	kind	type	lev	other inf
1	x	var	real	0	
2	y	var	real	0	
3	i	var	int	0	
4	k	var	int	0	
5	name	var	array	0	
6	M ₁	proc		0	
7	ind	para	int	1	
8	x	var	int	1	
9	P ₁	proc		1	
10	j	para	real	2	
11	P ₂	proc		2	
12	f	var	array	3	
13	test1	var	boolean	3	

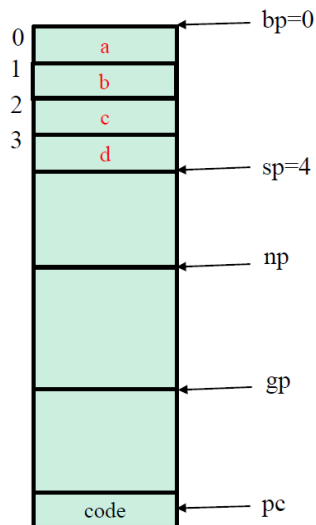
```
int a;
int b;
int c;
int d;
```

```
a = b + c + d;
b = a * a + b * b;
```

源程序

如何自动化?

编译器
对中间语言抽象机
及运行时存储组织
的设计

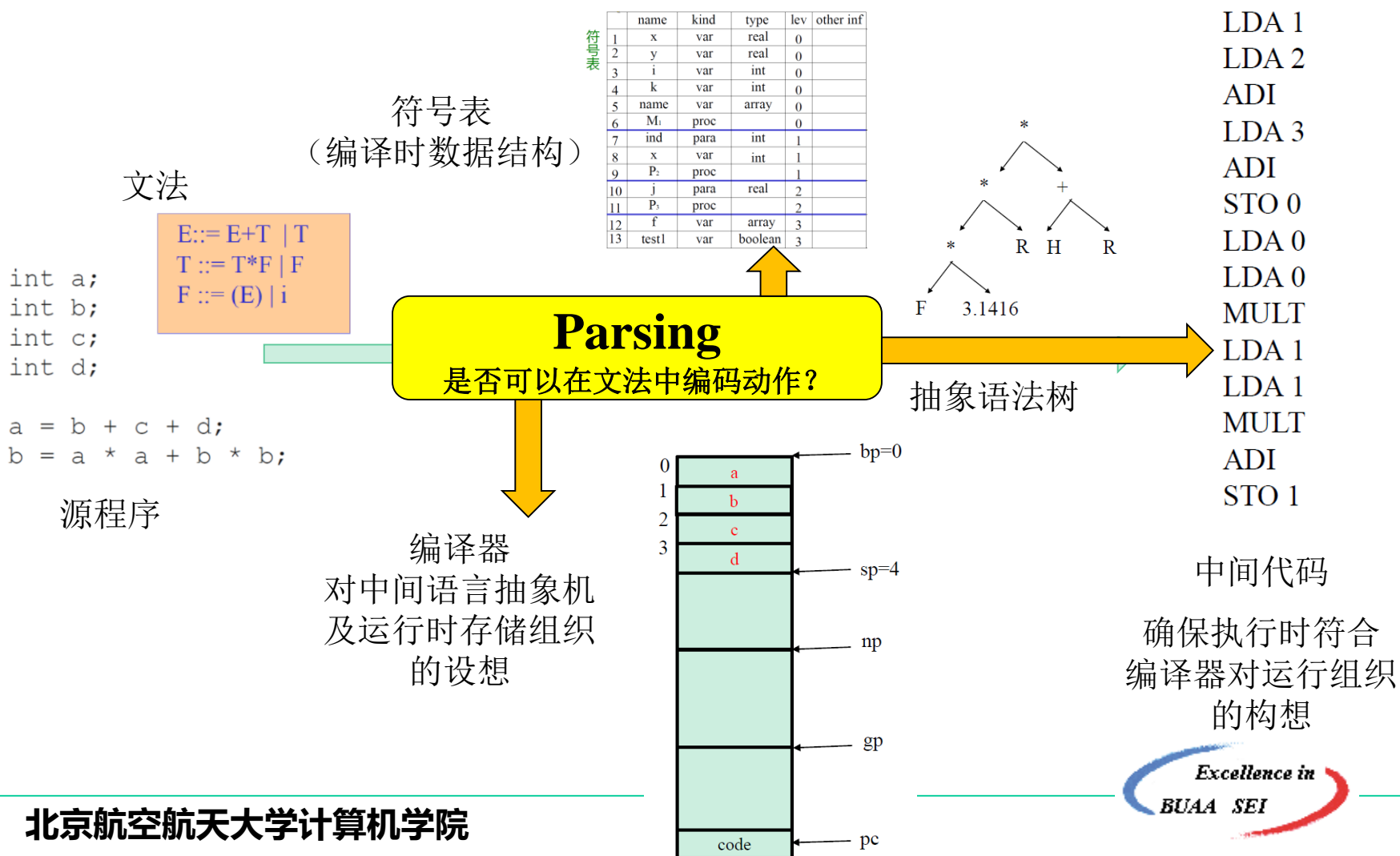


```
LDA 1
LDA 2
ADI
LDA 3
ADI
STO 0
LDA 0
LDA 0
MULT
LDA 1
LDA 1
MULT
ADI
STO 1
```

中间代码

确保执行时符合
编译器对运行组织
的构想

前两节课（符号表+运行时存储）：为代码生成做“需求分析”



第九章 语法制导翻译技术

- 翻译文法(TG)和语法制导翻译
- 属性翻译文法(ATG)
- 自顶向下语法制导翻译
 - 翻译文法的自顶向下语法制导翻译
 - 属性文法的自顶向下语法制导翻译

翻译文法

9.1 翻译文法和语法制导翻译

有上下无关文法 $G[E]$:

$$1. E \rightarrow E+T$$

$$4. T \rightarrow F$$

$$2. E \rightarrow T$$

$$5. F \rightarrow (E)$$

$$3. T \rightarrow T * F$$

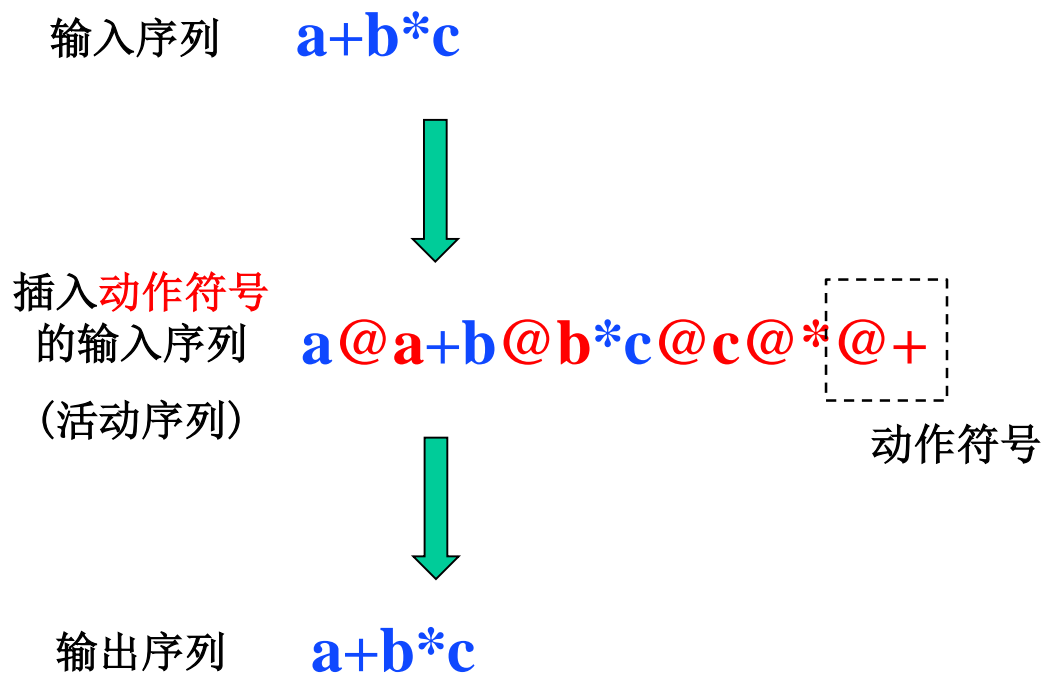
$$6. F \rightarrow i$$

此文法是一个中缀算术表达式文法

翻译的任务是: 中缀表达式 \Rightarrow 逆波兰表示

$$a+b*c \Rightarrow abc*+$$

假如翻译任务是要将中缀表达式简单变换为波兰后缀表示, 只需在上述文法中插入相应的动作符号。



1. $E \rightarrow E+T$
2. $E \rightarrow T$
3. $T \rightarrow T * F$

4. $T \rightarrow F$
5. $F \rightarrow (E)$
6. $F \rightarrow a|b|c$

输入序列 **a+b*c**

输入文法

$E \rightarrow E+T \rightarrow E+T * F \rightarrow E+T * c \rightarrow E+F * c \rightarrow E+b * c$
 $\rightarrow T+b * c \rightarrow F+b * c \rightarrow a+b * c$

插入动作符号
的输入序列
(活动序列)

a@a+b@b*c@c@*@+

动作符号

活动序列语言文法

1. $E \rightarrow E+T @+$
2. $E \rightarrow T$
3. $T \rightarrow T * F @*$
4. $T \rightarrow F$
5. $F \rightarrow (E)$
6. $F \rightarrow a@a | b@b | c@a$

输出序列 **a+b*c**

$E \rightarrow E+T @+ \rightarrow E+T * F @* @+ \rightarrow E+T * c @c @* @+ \rightarrow E+F * c @c @* @+ \rightarrow E+b @b * c @c @* @+ \rightarrow T+b @b * c @c @* @+ \rightarrow F+b @b * c @c @* @+ \rightarrow a@a+b @b * c @c @* @+$

1. $E \rightarrow E+T$
2. $E \rightarrow T$
3. $T \rightarrow T * F$

4. $T \rightarrow F$
5. $F \rightarrow (E)$
6. $F \rightarrow i$

1. $E \rightarrow E+T @+$
2. $E \rightarrow T$
3. $T \rightarrow T * F @*$
4. $T \rightarrow F$
5. $F \rightarrow (E)$
6. $F \rightarrow i @i$

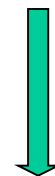
其中：

@+, @*, @i 为动作符号。@为动作符号标记，后面为字符串。
在本例中，其对应语义子程序的功能是要输出打印动作符号标记后面的字符串。

所以，产生式1： $E \rightarrow E+T @+$ 的语义是分析E, + 和T，输出+
产生式6： $F \rightarrow i @i$ 的语义是分析i，输出i

下面给出输入文法和翻译文法的概念：

- 输入文法：** 未插入动作符号时的文法。
由输入文法可以通过推导产生输入序列。
- 翻译文法：** 插入动作符号的文法。
由翻译文法可以通过推导产生活动序列。



{ 输入序列
动作序列

例: $(i+i) * i$

可以用输入文法推导:

$$E \Rightarrow T \Rightarrow T * F \Rightarrow F * F \Rightarrow (E) * F \Rightarrow (E + T) * F \\ \Rightarrow (i + i) * i$$

1. $E \rightarrow E + T$
2. $E \rightarrow T$
3. $T \rightarrow T * F$

4. $T \rightarrow F$
5. $F \rightarrow (E)$
6. $F \rightarrow i$

用相应的翻译文法推导, 可得:

$$E \Rightarrow T \\ \Rightarrow T * F @ * \\ \Rightarrow F * F @ * \\ \Rightarrow (E) * F @ * \\ \Rightarrow (E + T @ +) * F @ *$$

1. $E \rightarrow E + T @ +$
2. $E \rightarrow T$
3. $T \rightarrow T * F @ *$

4. $T \rightarrow F$
5. $F \rightarrow (E)$
6. $F \rightarrow i @ i$

$$\stackrel{*}{\Rightarrow} (i @ i + i @ i @ +) * i @ i @ *$$

活动序列：由翻译文法推导出的符号串，由终结符和动作符号组成。

- 从活动序列中，抽去动作符号，则得输入序列 $(i+i)*i$
- 从活动序列中，抽去输入序列，则得动作序列，执行动作序列，则完成翻译任务：

$$@i@i@+@i@* \Rightarrow ii+i*$$

定义9.1

翻译文法是上下文无关文法，其终结符号集由输入符号和动作符号组成。由翻译文法所产生的终结符号串称为**活动序列**。

语法制导的翻译

上例题中的翻译文法为：

$$G_T = (V_n, V_t, P, E)$$

$$V_n = \{E, T, F\}$$

$$V_t = \{i, +, *, (,), @ +, @ *, @ i\}$$

$$P = \{E \rightarrow E + T @ +, E \rightarrow T, T \rightarrow T * F @ *, T \rightarrow F, F \rightarrow (E), \\ F \rightarrow i @ i\}$$

符号串翻译文法：若插入文法中的动作符号对应的语义子程序是输出动作符号标记@后的字符串的文法。

语法制导翻译(Syntax Directed Translation)：按翻译文法进行的翻译。给定一输入符号串，根据翻译文法获得翻译该符号串的动作序列，并执行该序列所规定的动作的过程。

语法制导翻译的实现方法：

在文法的适当位置插入语义动作符号，当按文法分析到动作符号@时就调用相应的语义子程序，完成翻译任务。

翻译文法所定义的翻译是由输入序列和动作序列组成的对偶集。

如： $(i+i)*i$, $@i@i@+@i@*$ $\rightarrow ii + i*$

$i+i*i$ $@i@i@i@* @+$

因此，给定一个翻译文法，就给定了一个对偶集。

cān shù

带“属性”的翻译文法

9.2 属性翻译文法

在翻译文法的基础上，可以进一步定义**属性文法**，翻译文法中的符号，包括终结符、非终结符和动作符号均可带有属性，这样能更好的描述和实现编译过程。

属性可以分为两种：

综合属性

继承属性

9.2.1 综合属性

基本操作数带有属性的表达式文法G[E]

$$1. E \rightarrow E + F$$

$$2. E \rightarrow T$$

$$3. T \rightarrow T * F$$

$$4. T \rightarrow F$$

$$5. F \rightarrow (E)$$

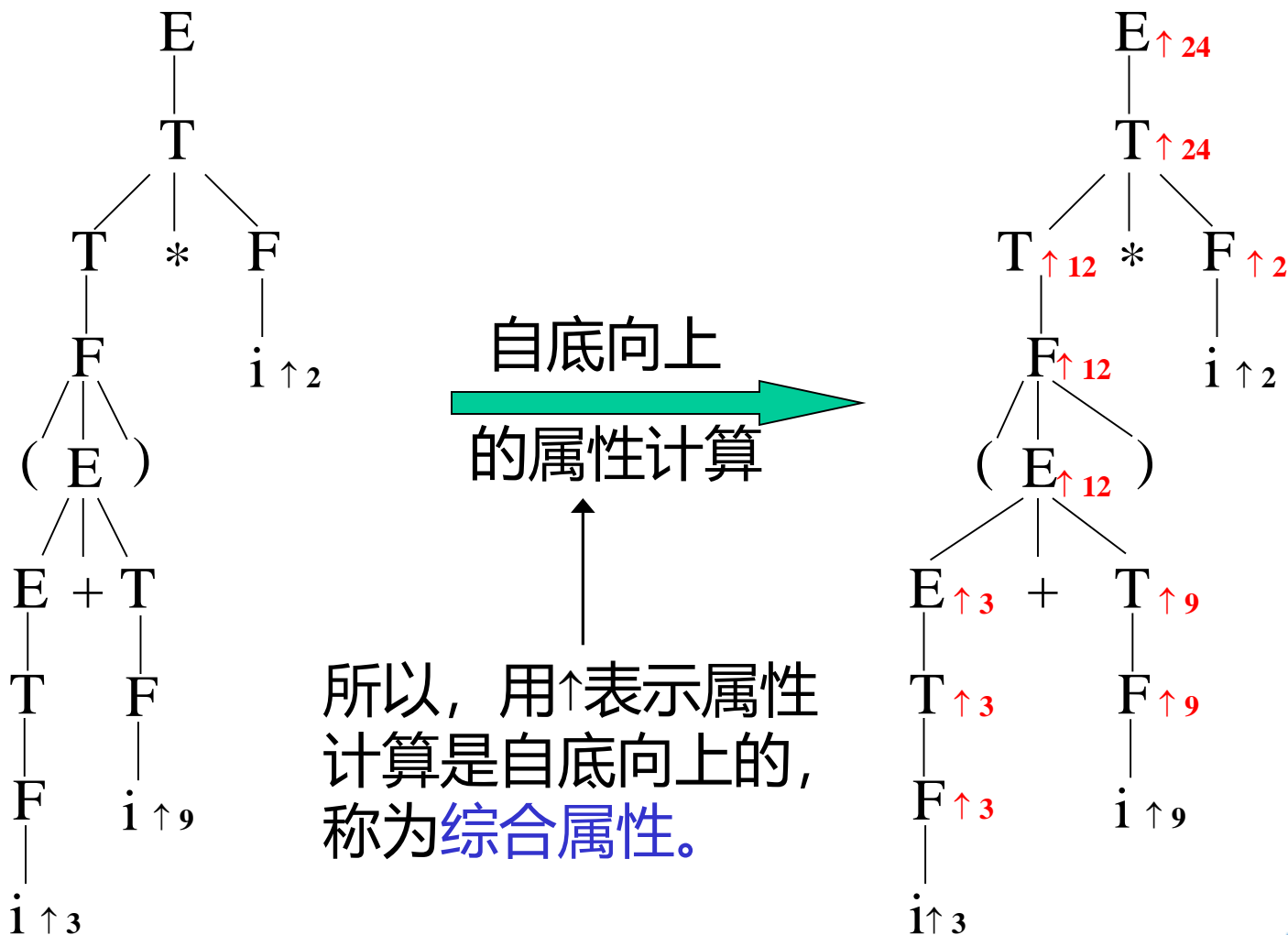
$$6. F \rightarrow i \uparrow c$$

其中 $\uparrow c$ 是综合属性符号, \uparrow 为综合属性标记, c 为属性变量或者属性值。

此文法能够产生如下的输入序列:

$$(i \uparrow_3 + i \uparrow_9) * i \uparrow_2$$

根据给定的文法，可写出该输入序列的语法树



为了形式地表示上述表达式的属性求值过程，可以改写上述文法：

产生式

1. $E \uparrow_{p4} \rightarrow E \uparrow_{q5} + T \uparrow_{r2}$
2. $E \uparrow_{p3} \rightarrow T \uparrow_{q4}$
3. $T \uparrow_{p2} \rightarrow T \uparrow_{q3} * F \uparrow_{r1}$
4. $T \uparrow_{p2} \rightarrow F \uparrow_{q2}$
5. $F \uparrow_{p1} \rightarrow (E \uparrow_{q1})$
6. $F \uparrow_{p1} \rightarrow i \uparrow_{q1}$

求值规则

- $$\begin{aligned} p_4 &:= q_5 + r_2 ; \\ p_3 &:= q_4 ; \\ p_2 &:= q_3 * r_1 ; \\ p_2 &:= q_2 ; \\ p_1 &:= q_1 ; \\ p_1 &:= q_1 ; \end{aligned}$$

说明：

- p, q, r 为属性变量名。
- 属性变量名局部于每个产生式，也可使用不同的名字。

- 求值规则：综合属性是自右向左，自底向上求值。

9.2.2 继承属性

考虑下列文法：G[<说明>]:

1. <说明> \rightarrow Type id <变量表>
2. <变量表> \rightarrow , id <变量表>
3. <变量表> $\rightarrow \epsilon$

其中

Type: 类型名 (值: int, real, bool等)

id: 变量名 (值: 指向该变量符号表项的指针)

上述文法所产生的语句: int A,BC

该文法的翻译任务: 将声明的变量填入符号表

完成该工作的动作符号: @set_table

int A, BC

符号表
A 整型
BC 整型

翻译文法:

1. $\langle \text{说明} \rangle \rightarrow \text{Type id @set_table } \langle \text{变量表} \rangle$
2. $\langle \text{变量表} \rangle \rightarrow , \text{id @set_table } \langle \text{变量表} \rangle$
3. $\langle \text{变量表} \rangle \rightarrow \varepsilon$

填表时需要的信息：类型，名字，以及填的位置如何得到？

类型、名字在词法分析时得到，可设两个综合属性。

填写位置：可以维护一个运行时组织分配的全局指针。

Type $\uparrow t$

t中放类型值

id $\uparrow n$

n中放变量名

填表动作符号也可带有属性：

@set_table \downarrow_{t_1, n_1}

\downarrow_{t_1, n_1}

可从前面得到，所以称为继承属性，
继承前面的值

<变量表> \downarrow_{t_2}

\downarrow_{t_2}

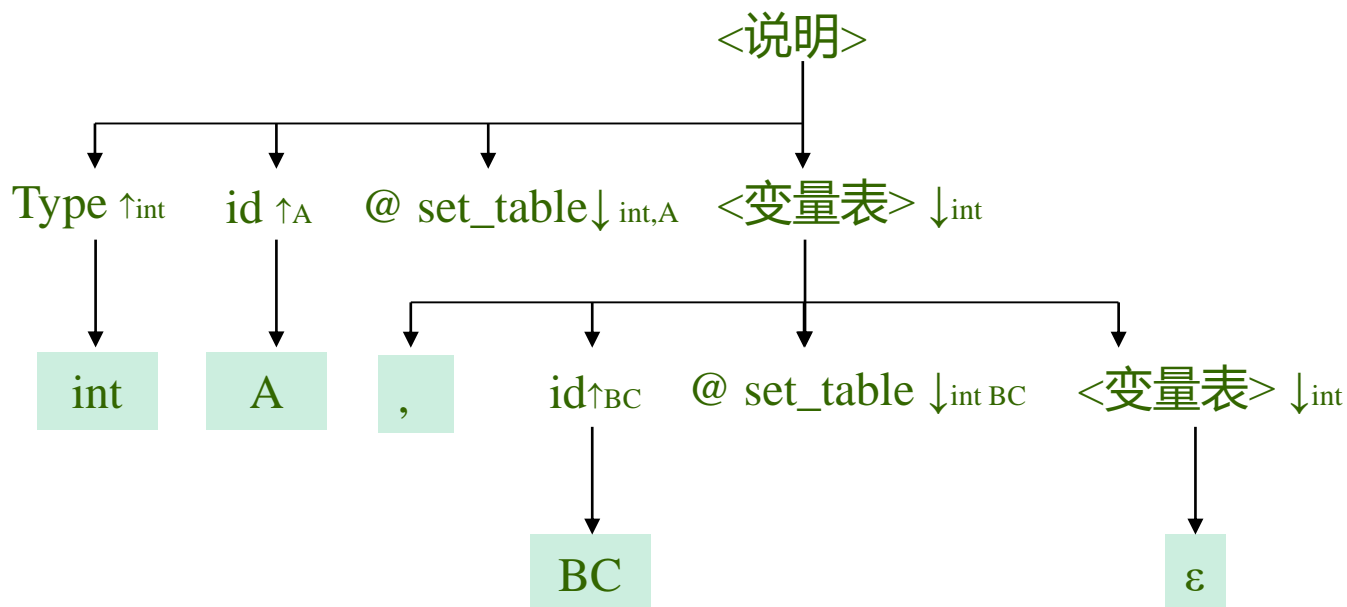
同上

属性翻译文法：

1. <说明> \rightarrow Type $\uparrow t$ id $\uparrow n$ @set_table \downarrow_{t_1, n_1} <变量表> \downarrow_{t_2} $t_2, t_1 := t; \quad n_1 := n;$
2. <变量表> $\downarrow_{t_2} \rightarrow ,$ id $\uparrow n$ @set_table \downarrow_{t_1, n_1} <变量表> \downarrow_{t_3} $t_3, t_1 := t_2; \quad n_1 := n;$
3. <变量表> $\downarrow_{t_2} \rightarrow \epsilon$

1. $\langle \text{说明} \rangle \rightarrow \text{Type} \uparrow_t \text{ id} \uparrow_n @ \text{set_table} \downarrow_{t1, n1} \langle \text{变量表} \rangle \downarrow_{t2} \quad t_2, t_1 := t; \quad n_1 := n;$
2. $\langle \text{变量表} \rangle \downarrow_{t2} \rightarrow , \text{ id} \uparrow_n @ \text{set_table} \downarrow_{t1, n1} \langle \text{变量表} \rangle \downarrow_{t3} \quad t_3, t_1 := t_2; \quad n_1 := n;$
3. $\langle \text{变量表} \rangle \downarrow_{t2} \rightarrow \varepsilon$

例: $\text{int } A, BC \Rightarrow \text{Type} \uparrow_{\text{int}} \text{ id} \uparrow_A, \text{ id} \uparrow_{BC}$



符号表
A 整型
BC 整型

继承属性求值：自左向右，自顶向下
综合属性求值：自右向左，自底向上

int A, BC 在填表中的地址处理:

$\langle \text{说明} \rangle \Rightarrow \text{Type} \uparrow_t \text{id} \uparrow_{n1} @ \text{set_table} \downarrow_{t,n1} \langle \text{变量表} \rangle \downarrow_t$
 $\stackrel{+}{\Rightarrow} \text{Type} \uparrow_t \text{id} \uparrow_{n1} @ \text{set_table} \downarrow_{t,n1}$
 $\quad , \text{id} \uparrow_{n2} @ \text{set_table} \downarrow_{t,n2}$

符号表

p \longrightarrow

A	int
BC	int

L-属性翻译文法 (L-TAG)

9.2.3 (1) L-属性翻译文法 (L-ATG)

这是属性翻译文法中较简单的一种。其输入**文法要求是LL(1)文法**，可用自顶向下分析构造分析器。在分析过程中可进行属性求值。

定义9.2:

L-属性翻译文法是带有下列说明的翻译文法:

- 1. 文法中的终结符，非终结符及动作符号都带有属性，且每个属性都有一个值域。**
- 2. 非终结符及动作符号的属性可分为继承属性和综合属性。**
- 3. 开始符号的继承属性具有指定的初始值。**
- 4. 输入符号（终结符号）的每个综合属性具有指定的初始值。**
- 5. 属性的求值规则:**

属性的求值规则：

体现自顶向下，自左向右的求值特性。

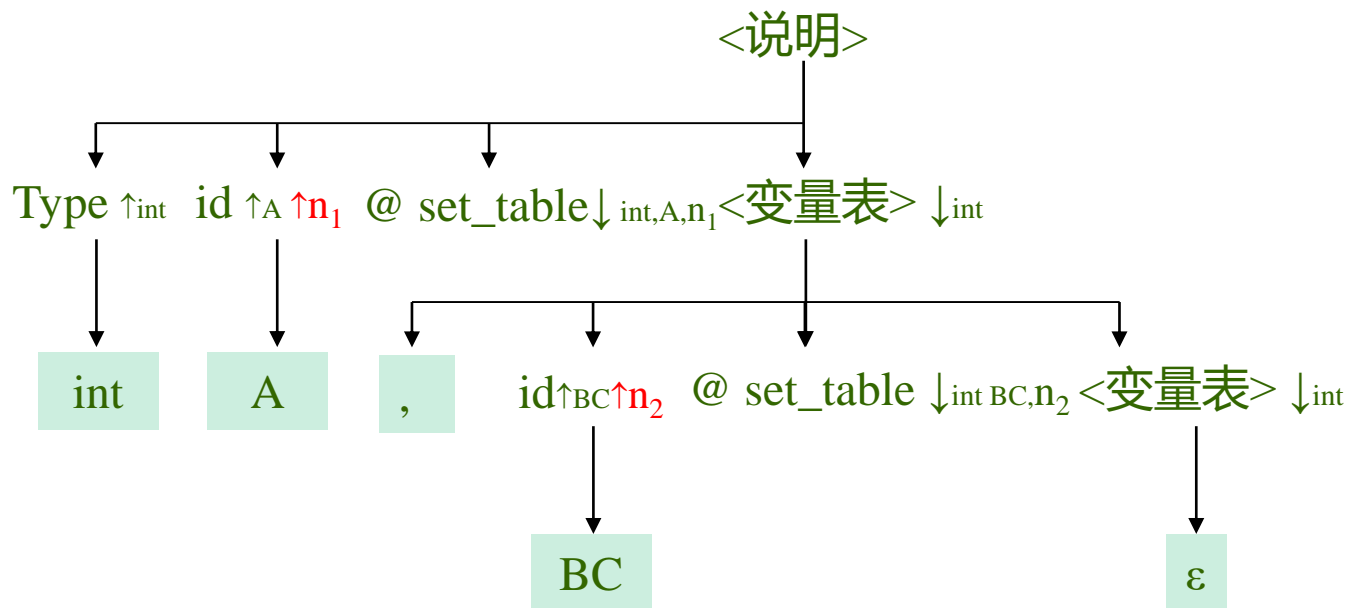
继承属性：

- (1) 产生式左部非终结符号的继承属性值，取前面产生式右部该符号已有的继承属性值。
- (2) 产生式右部符号的继承属性值，用该产生式左部符号的继承属性或出现在该符号左部的符号的属性值进行计算。

综合属性：

- (1) 产生式右部非终结符号的综合属性值，取其下部产生式左部同名非终结符号的综合属性值。
- (2) 产生式左部非终结符号的综合属性值，用该产生式左部符号的继承属性或某个右部符号的属性进行计算。
- (3) 动作符号的综合属性用该符号的继承属性或某个右部符号的属性进行计算。

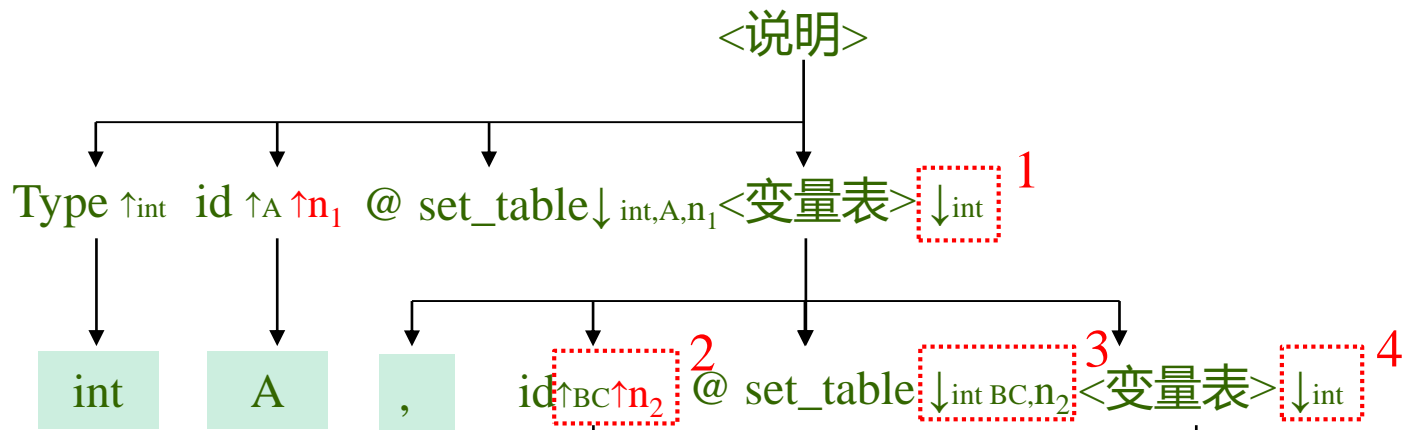
适合在自顶向下分析过程中求值



例: $A \rightarrow BC$

求值顺序:

- 1) A的继承属性 \downarrow (若A为开始符号, 则有指定值, 否则由上面产生式右部符号的继承属性求得)
- 2) B的继承属性 \downarrow (由A的继承属性求得)
- 3) B的综合属性 \uparrow (由下面产生式中左部符号为B的综合属性求得)



(2) 产生式左部非终结符号的综合属性值，用该产生式左部符号的继承属性或某个右部符号的属性进行计算。

例：A → BC
求值顺序：

- | | |
|------------------------|---|
| 1) A的继承属性 ↓
产生式右部符号 | 4) C的继承属性 ↓ (由A的继承属性和B的属性求得) |
| 2) B的继承属性 ↓ | 5) C的综合属性 ↑ (由下面产生式中左部符号为C的综合属性求得) |
| 3) B的综合属性 ↑
求得) | 6) A的综合属性 ↑ (由前述(2)，即A的继承属性或产生式某右部符号属性计算) |

简单赋值形式的L-TAG (SL-TAG)

(2) 简单赋值形式的L_属性翻译文法(SL-ATG)

- 一般属性值计算: $x := f(y, z)$
SL-ATG属性值计算: $x := \text{某符号的属性值或常量}$
例 $x := y, \quad x, y, z := 17$ —— 称为复写规则

为了实现上的方便，常希望文法符号的属性求值规则为上述简单形式的。为此，对现有的L-ATG的定义做一点改变，从而形成一个称为简单赋值形式的L-ATG。

• **定义9.4** 一个L-ATG被定义为简单赋值形式的(SL-ATG), 当且仅当满足如下条件:

1. **产生式右部符号的继承属性 \downarrow** 是一个常量, 它等于左部符号的继承属性值或等于出现在所给符号左边符号的一个综合属性值。

2. **产生式左部非终结符号的综合属性 \uparrow** 是一个常量, 它等于左部符号的继承属性值或等于右部符号的综合属性值。

因此, 一个简单赋值形式的L-ATG除动作符号外, 其余符号的属性求值规则其右部是属性或是常量。

- L-ATG \Rightarrow SL-ATG

给定一个L-ATG，如何找一个等价的赋值形式的L-ATG？

考虑产生式:

$$\langle A \rangle \rightarrow a \uparrow_R \langle B \rangle \uparrow_S \langle C \rangle \downarrow_I, \quad I := f(R, S)$$

显然: 该属性求值规则不是简单赋值形式的，因为它需要对f求值。

思路: 把求值函数变成一个动作@f。

$$\langle A \rangle \rightarrow a \uparrow_R \langle B \rangle \uparrow_S \langle C \rangle \downarrow_I, \quad I := f(R, S)$$

第一步：设动作符号 “@ f” 表示函数f求值，该动作符号有两个继承属性和一个综合属性。

$$@f \downarrow_{I_1, I_2} \uparrow_{S_1} \quad \text{且} \quad S_1 := f(I_1, I_2)$$

第二步：修改产生式

1. 插入 “@ f” (在适当位置)
2. 引进新的复写规则 (将R, S 赋给I₁和I₂, f值赋给S₁)
3. 删去原有包含f 的规则

$$\langle A \rangle \rightarrow a \uparrow_R \langle B \rangle \uparrow_S @f \downarrow_{I_1, I_2} \uparrow_{S_1} \langle C \rangle \downarrow_I,$$

$$I_1 := R, \quad I_2 := S, \quad S_1 := f(I_1, I_2), \quad I := S_1$$

$$\langle A \rangle \rightarrow @ f_{\downarrow I_1, I_2 \uparrow S_1} a \uparrow_R \langle B \rangle \uparrow_S \langle C \rangle \downarrow_I,$$

$$I_1 := R, \quad I_2 := S, \quad S_1 := f(I_1, I_2), \quad I := S_1$$

$$\langle A \rangle \rightarrow a \uparrow_R @ f_{\downarrow I_1, I_2 \uparrow S_1} \langle B \rangle \uparrow_S \langle C \rangle \downarrow_I,$$

$$I_1 := R, \quad I_2 := S, \quad S_1 := f(I_1, I_2), \quad I := S_1$$

$$\langle A \rangle \rightarrow a \uparrow_R \langle B \rangle \uparrow_S @ f_{\downarrow I_1, I_2 \uparrow S_1} \langle C \rangle \downarrow_I,$$

$$I_1 := R, \quad I_2 := S, \quad S_1 := f(I_1, I_2), \quad I := S_1.$$

$$\langle A \rangle \rightarrow a \uparrow_R \langle B \rangle \uparrow_S \langle C \rangle \downarrow_I @ f_{\downarrow I_1, I_2 \uparrow S_1},$$

$$I_1 := R, \quad I_2 := S, \quad S_1 := f(I_1, I_2), \quad I := S_1..$$

什么是合适位置？

(2) 右部符号的继承属性值，用该产生式左部符号的继承属性或出现在该符号左部的符号的属性值进行计算。

(3) 动作符号的综合属性用该符号的继承属性或某个右部符号的属性进行计算。

$$\langle A \rangle \rightarrow a \uparrow_R \langle B \rangle \uparrow_S \langle C \rangle \downarrow_I, \quad I := f(R, S)$$

$$\langle A \rangle \rightarrow a \uparrow_R \langle B \rangle \uparrow_S @ f \downarrow_{I_1, I_2} \uparrow_{S_1} \langle C \rangle \downarrow_I,$$

$$I_1 := R, \quad I_2 := S, \quad S_1 := f(I_1, I_2), \quad I := S_1.$$

该文法是简单赋值形式的L-ATG.

注意： 无参函数过程作为常数处理，如

$$\langle A \rangle \rightarrow \langle B \rangle \uparrow_x \langle C \rangle \uparrow_y \quad x, y := \text{NEWT}$$

实现：自顶向下翻译（递归下降）

9.3 自顶向下语法制导翻译

9.3.1 翻译文法的自顶向下翻译 ——递归下降翻译器

9.3.2 属性翻译文法的自顶向下翻译的实现 ——递归下降属性翻译器

9.3.1 翻译文法的自顶向下翻译——递归下降翻译器

按翻译要求，在文法中插入语义动作符号，在分析过程中调用相应的语义处理程序，完成翻译任务。

例：输入文法

1. $\langle S \rangle \longrightarrow a \langle A \rangle \langle S \rangle$
2. $\langle S \rangle \longrightarrow b$
3. $\langle A \rangle \longrightarrow c \langle A \rangle \langle S \rangle b$
4. $\langle A \rangle \longrightarrow \epsilon$

翻译文法（符号串翻译文法）

- $$\begin{aligned} \langle S \rangle &\longrightarrow a \langle A \rangle @ x \langle S \rangle \\ \langle S \rangle &\longrightarrow b @ z \\ \langle A \rangle &\longrightarrow c @ y \langle A \rangle \langle S \rangle @ v b \\ \langle A \rangle &\longrightarrow @ w \end{aligned}$$

1. $\langle S \rangle \rightarrow a \langle A \rangle \langle S \rangle$
2. $\langle S \rangle \rightarrow b$
3. $\langle A \rangle \rightarrow c \langle A \rangle \langle S \rangle b$
4. $\langle A \rangle \rightarrow \epsilon$

- $$\begin{aligned} \langle S \rangle &\rightarrow a \langle A \rangle @ x \langle S \rangle \\ \langle S \rangle &\rightarrow b @ z \\ \langle A \rangle &\rightarrow c @ y \langle A \rangle \langle S \rangle @ v b \\ \langle A \rangle &\rightarrow @ w \end{aligned}$$

主程序

```

if CLASS = a or b      then
    PROCS;
if CLASS ≠ 右界符    then
    ERROR;
ACCEPT;

```

过程 PROCS

```

case CLASS of
a : P1;
b : P2;
其它: ERROR;
end of case;

```

主程序

```

if CLASS = a or b      then
    PROCS;
if CLASS ≠ #          then
    ERROR;
ACCEPT;

```

过程 PROCS

```

case CLASS of
a : P1;
b : P2;
其它: ERROR;
end of case.

```

1. $\langle S \rangle \rightarrow a \langle A \rangle \langle S \rangle$
2. $\langle S \rangle \rightarrow b$
3. $\langle A \rangle \rightarrow c \langle A \rangle \langle S \rangle b$
4. $\langle A \rangle \rightarrow \epsilon$

- $\langle S \rangle \rightarrow a \langle A \rangle @ x \langle S \rangle$
- $\langle S \rangle \rightarrow b @ z$
- $\langle A \rangle \rightarrow c @ y \langle A \rangle \langle S \rangle @ v b$
- $\langle A \rangle \rightarrow @ w$

P₁: / *产生式1的代码* /

NEXTSYM;

PROCA;

PROCS;

RETURN;

P₂: / *产生式2的代码* /

NEXTSYM;

RETURN;

过程 PROCA

...

P₁: / *产生式1的代码* /

NEXTSYM;

PROCA;

OUT(x);

PROCS;

RETURN;

P₂: / *产生式2的代码* /

NEXTSYM;

OUT(z);

RETURN;

过程 PROCA

...

9.3.2 属性文法自顶向下翻译的实现——递归下降翻译器

方法：

- 对于每个非终结符号都编写一个翻译子程序（过程）。根据该非终结符号具有的属性数目，设置相应的参数。

继承属性：声明为赋值形参

综合属性：声明为变量形参

$U_{\downarrow x, \uparrow y} \longrightarrow \dots$

Procedure U(x, &y);

x—赋值形参

y—变量形参

- 过程调用语句的实参：

继承属性：继承属性值

综合属性：属性变量名（传地址，返回时有值）

- 关于属性名的约定：

1) 产生式左部的同名非终结符使用相同的属性名。

(递归下降分析法所必须)

$$\begin{array}{ll} \langle L \rangle \uparrow_a \downarrow_b \rightarrow e \downarrow_I \langle R \rangle \downarrow_J & \langle L \rangle \uparrow_x \downarrow_y \rightarrow e \downarrow_I \langle R \rangle \downarrow_J \\ \langle L \rangle \uparrow_x \downarrow_y \rightarrow \langle H \rangle \downarrow_z \uparrow_w & \langle L \rangle \uparrow_x \downarrow_y \rightarrow \langle H \rangle \downarrow_z \uparrow_w \end{array}$$

2) 具有相同值的属性取相同的属性名。

具有简单赋值形式的属性变量名取相同的属性名，可删去属性求值规则。

$$\langle S \rangle \rightarrow I \uparrow_a \langle B \rangle \downarrow_b \langle C \rangle \downarrow_c \quad b, c := a$$

$$\langle S \rangle \rightarrow I \uparrow_x \langle B \rangle \downarrow_x \langle C \rangle \downarrow_x$$

构造属性文法的递归下降翻译器：

例：有如下属性翻译文法 $G[< S >]$

1. $< S >_{\downarrow R_1} \rightarrow a \uparrow_{T_1} < A > \uparrow_{Q_1} @ x \downarrow_{T_2, R_2} < S > \downarrow_{Q_2}$
 $R_2 := R_1$
 $T_2 := T_1$
 $Q_2 := Q_1$
2. $< S >_{\downarrow R_1} \rightarrow b @ z \downarrow_{R_2}, \quad R_2 := R_1$
3. $< A > \uparrow_P \rightarrow C \uparrow_{U_1} @ y \downarrow_{U_2} < A > \uparrow_Q < S > \downarrow_z @ v \downarrow_P b$
 $U_2 := U_1, P := Q + U_1, z := U_1 - 3$
4. $< A > \uparrow_P \rightarrow @ w \quad P := 8$

对简单赋值形式的属性变量取相同的属性名，其求值规则可以删去。开始符号的继承属性 $R_1=7$ 。

全局变量和过程声明:

```
CLASS;           /* 存放单词类别码 */  
TOKEN;          /* 存放单词值 */  
NEXTSYM;        /* 词法分析程序, 每调用一次单词类别码  $\Rightarrow$  CLASS,  
                单词值  $\Rightarrow$  TOKEN, 该符号指针指向下一个单词 */
```

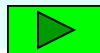
主程序:

```
NEXTSYM;  
PROCS(7);        // 开始符号的继承属性  $R_1=7$   
if CLASS  $\neq$  右界符 then ERROR;
```

```
ACCEPT
```

过程

```
PROCS(R)  
R;              /* 值形参声明 */
```



1. $\langle S \rangle_{\downarrow R} \longrightarrow a \uparrow_T \langle A \rangle \uparrow_Q @ X_{\downarrow T, R} \langle S \rangle_{\downarrow Q}$
2. $\langle S \rangle_{\downarrow R} \longrightarrow b @ Z_{\downarrow R}$

case CLASS of // First(S) = {a, b}

a: P_1 ;

b: P_2 ;

其它: ERROR;

end of case;

P_1 : /* 产生式1的代码 */

T , Q ; /* 局部变量声明 */

T := TOKEN; /* 单词值赋给终结符的综合属性 */

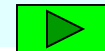
NEXTSYM;

PROCA(Q)

OUT($X_{\downarrow T, R}$) ;

PROCS(Q)

RETURN;



2. $\langle S \rangle_{\downarrow R} \longrightarrow b @ z_{\downarrow R}$
3. $\langle A \rangle_{\uparrow P} \longrightarrow C_{\uparrow u} @ y_{\downarrow u} \langle A \rangle_{\uparrow Q} \langle S \rangle_{\downarrow z} @ v_{\downarrow P} b$
4. $\langle A \rangle_{\uparrow P} \longrightarrow @ w \quad P:=8$

P_2 : / * 产生式2的代码 * /

NEXTSYM;

OUT($Z_{\downarrow R}$);

RETURN;

过程 PROC A(P)

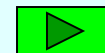
P; /*变量形参声明*/

case CLASS of

 C : p3

 其它: p4

end of case;



3. $\langle A \rangle \uparrow_P \longrightarrow C \uparrow_u @ y \downarrow_u \langle A \rangle \uparrow_Q \langle S \rangle \downarrow_z @ V \downarrow_P b$
 $P := Q + U, z := U - 3$

P3:

U , Q , Z ;

/*局部变量声明*/

U := TOKEN;

插在U已知, 使用Z之前

NEXTSYM;

Z := U - 3 ;

OUT($y \downarrow_u$);

返回时有值

PROC A(Q);

插在Q,U已知, 使用P之前。

P := Q+U;

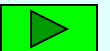
PROC S(Z);

OUT($V \downarrow_P$);

if CLASS \neq b then ERROR;

NEXTSYM;

RETURN;



4. $\langle A \rangle_{\uparrow P} \longrightarrow @ w \quad P := 8$

P4:

$P := 8;$

OUT(w);

RETURN;

例子：算术表达式生成中间代码

一个例子

例：构造将算术表达式翻译成四元式的属性翻译文法，并写出递归下降分析程序。由该属性翻译文法来描述翻译过程。

翻译的输入：算术表达式 $a + b$

翻译的输出：四元式 ADD, P_a, P_b, P_r
其中 P_a, P_b, P_r 为变量 a, b 和结果单元的地址。

表达式： $(a + b) * c$

输入： $(Id \uparrow_1 + Id \uparrow_2) * Id \uparrow_4$
 Id 由词法分析程序返回，
 \uparrow_1综合属性，变量在数据区地址。

输出： $ADD, 1, 2, 3$
 $MULT, 3, 4, 5$

数据区：

1	a
2	b
3	部分结果
4	c
5	部分结果

(1) 翻译文法设计:

$E \rightarrow E + T \text{ @ADD}$

$T \rightarrow F$

$E \rightarrow T$

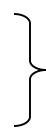
$F \rightarrow (E)$

$T \rightarrow T * F \text{ @MULT}$

$F \rightarrow \text{Id}$

@ADD为输出ADD四元式的动作符号

@MULT为输出MULT四元式的动作符号



对应于完成翻译
的语义动作程序

在文法中的插入位置: 在分别处理完成两个操作数之后

输入序列: $(a+b)*c$

翻译文法产生的活动序列: $(a+b \text{ @ADD}) * c \text{ @MULT}$

动作符号序列: @ADD @MULT

反映生成四元式的顺序, 语法分析过程中语义程序的调用顺序。

(2) 属性翻译文法的设计

- 输入符号（操作数）有一个综合属性，它是该符号在数据区的地址。
- 每个非终结符有一个综合属性，该属性是由它产生的代表该子表达式在数据区的地址。
- 动作符号有三个继承属性，它们分别是左右操作数和运算结果在数据区地址。

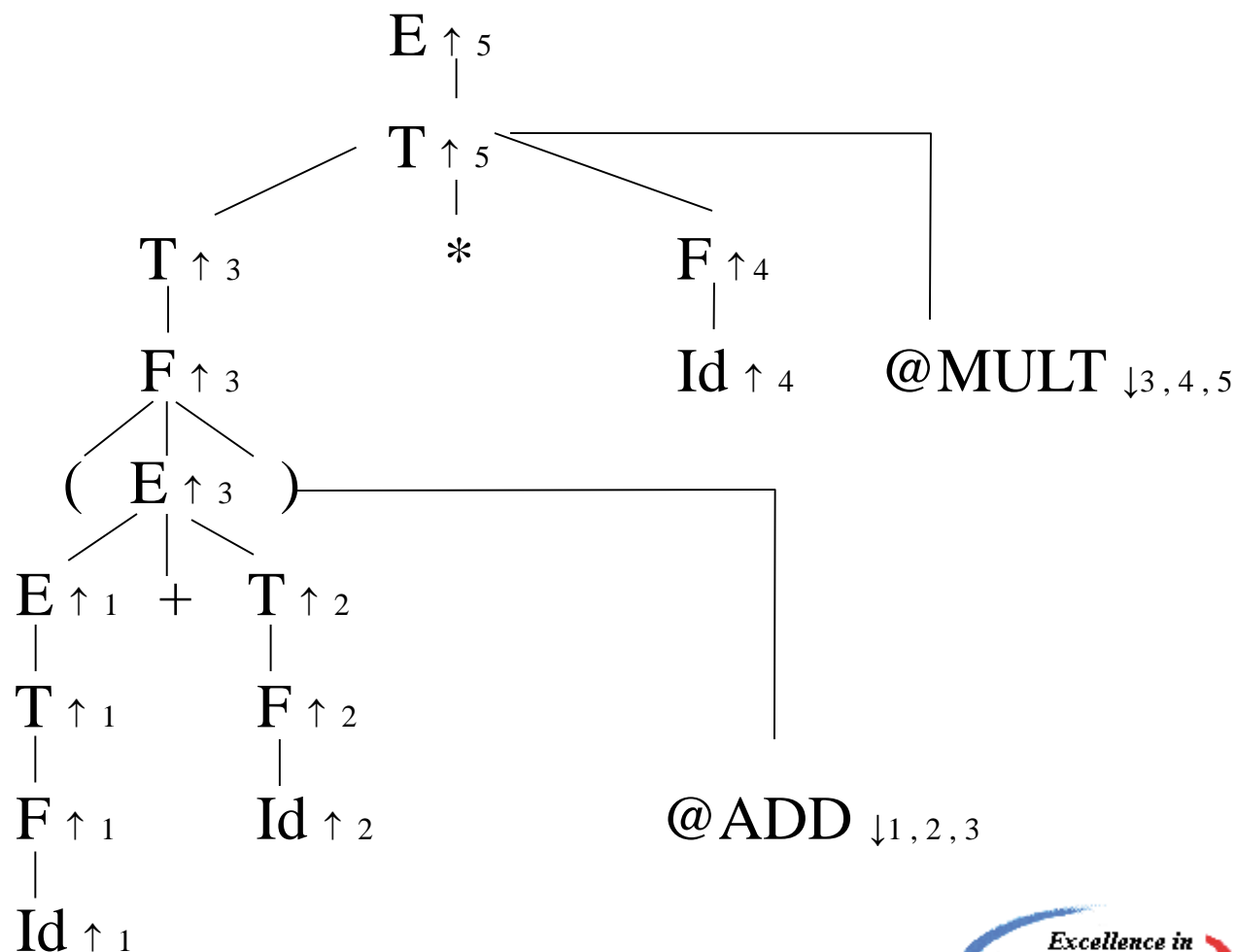
这样可得表达式的属性翻译文法
——可将中缀表达式翻译成四元式

- | | | | |
|--|-----------------------|----------|----------|
| 1. $E \uparrow_x \rightarrow E \uparrow_q + T \uparrow_r @ADD_{\downarrow y, z, p}$ | $x, p := \text{NEWT}$ | $y := q$ | $z := r$ |
| 2. $E \uparrow_x \rightarrow T \uparrow_p$ | $x := p$ | | |
| 3. $T \uparrow_x \rightarrow T \uparrow_q * F \uparrow_r @MULT_{\downarrow y, z, p}$ | $x, p := \text{NEWT}$ | $y := q$ | $z := r$ |
| 4. $T \uparrow_x \rightarrow F \uparrow_p$ | $x := p$ | | |
| 5. $F \uparrow_x \rightarrow (E \uparrow_p)$ | $x := p$ | | |
| 6. $F \uparrow_x \rightarrow \text{Id} \uparrow_p$ | $x := p$ | | |

说明:

Id的综合属性p是数据区地址, NEWT为系统过程, 返回数据区地址。

反映属性求值的语法树：



让我们从头开始：原始文法

1. $E \rightarrow E + E$

2. $E \rightarrow E * E$

3. $E \rightarrow (E)$

3. $E \rightarrow i$

二义性文法?



1. $E \rightarrow E + T \mid T$

2. $T \rightarrow T * F \mid F$

3. $F \rightarrow (E)$

3. $F \rightarrow i$

无二义性文法

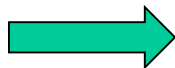
1. $E \rightarrow E + T \mid T$

2. $T \rightarrow T * F \mid F$

3. $F \rightarrow (E)$

3. $F \rightarrow i$

左递归?



1. $E \rightarrow TE'$

2. $E' \rightarrow +TE' \mid \varepsilon$

3. $T \rightarrow FT'$

4. $T' \rightarrow *FT' \mid \varepsilon$

5. $F \rightarrow (E)$

6. $F \rightarrow i$

规则一 (提因子)

若: $U ::= xy | xw | \dots | xz$

则可改写为: $U ::= x(y | w | \dots | z)$

规则三

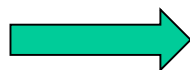
若: $P ::= P\alpha \mid \beta$

则可改写为: $P ::= \beta P'$

$P' ::= \alpha P' \mid \varepsilon$

右递归的无二义性文法

增加?

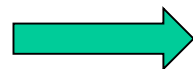


1. $E \rightarrow TE'$
2. $E' \rightarrow +TE' \mid \varepsilon$
3. $T \rightarrow FT'$
4. $T' \rightarrow *FT' \mid \varepsilon$
5. $F \rightarrow (E)$
6. $F \rightarrow i$

1. $E_{\uparrow x} \rightarrow T_{\uparrow p} E'_{\uparrow q} @+_{\downarrow p,q,r \uparrow s}$
2. $E'_{\uparrow x} \rightarrow + T_{\uparrow p} E'_{\uparrow q} @+_{\downarrow p,q,r \uparrow s}$
3. $E'_{\uparrow x} \rightarrow \varepsilon \quad (x=0)$
4. $T_{\uparrow x} \rightarrow FT' @*_{\downarrow p,q,r \uparrow s}$
5. $T'_{\uparrow x} \rightarrow *F_{\uparrow p} T'_{\uparrow q} @*_{\downarrow p,q,r \uparrow s}$
6. $T'_{\uparrow x} \rightarrow \varepsilon \quad (x=1)$
7. $F_{\uparrow x} \rightarrow (E_{\uparrow p}) \quad (x=p)$
8. $F_{\uparrow x} \rightarrow i_{\uparrow p} \quad (x=p)$

简单赋值的属性翻译文法 (SL-ATG)

1. $E_{\uparrow x} \rightarrow T_{\uparrow p} E'_{\uparrow q} @+_{\downarrow p,q,r \uparrow s}$
2. $E'_{\uparrow x} \rightarrow + T_{\uparrow p} E'_{\uparrow q} @+_{\downarrow p,q,r \uparrow s}$
3. $E'_{\uparrow x} \rightarrow \varepsilon \quad (x=0)$
4. $T_{\uparrow x} \rightarrow FT' @*_{\downarrow p,q,r \uparrow s}$
5. $T'_{\uparrow x} \rightarrow *F_{\uparrow p} T'_{\uparrow q} @*_{\downarrow p,q,r \uparrow s}$
6. $T'_{\uparrow x} \rightarrow \varepsilon \quad (x=1)$
7. $F_{\uparrow x} \rightarrow (E_{\uparrow p}) \quad (x=p)$
8. $F_{\uparrow x} \rightarrow i_{\uparrow p} \quad (x=p)$



递归下降的
属性翻译器

小结:

本章介绍了语法制导翻译的概念和技术，是在抽象层次上讲的。

第十章对过程语言的编译就是采用该法。

翻译文法：在输入文法中插入语义动作符号（完成翻译任务的语义子程序）

属性翻译文法：文法符号（包括动作符号）可带有属性，并定义相应的属性求值规则，就成为属性翻译文法。比翻译文法能更细地描述翻译过程。（属性有综合属性和继承属性之分）

程序语言的属性翻译文法都是L-属性的，一般无求值规则。

语义动作程序如何设计在后面介绍

$@ADD \downarrow y, z, p \Rightarrow \text{fprintf}(\text{objfile}, \text{"ADD \%d \%d \%d \n"}, y, z, p)$

(3) 写递归下降翻译程序 (实践课程)

作业： P166 1 (前缀式), 2, 3, 4, 5

P180 2, 3 (简单赋值形式重写)

谢谢!