

语法分析

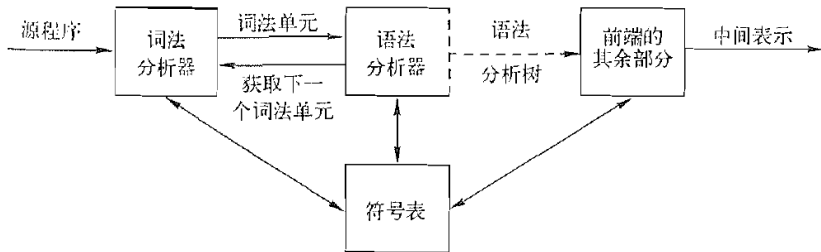
魏恒峰

hfwei@nju.edu.cn

2020 年 11 月 29 日



输入：词法单元流 & 语言的语法规则



输出：语法分析树 (Parse Tree)

语法分析举例

$\langle \text{Stmt} \rangle \rightarrow \langle \text{Id} \rangle = \langle \text{Expr} \rangle ;$
$\langle \text{Stmt} \rangle \rightarrow \{ \langle \text{StmtList} \rangle \}$
$\langle \text{Stmt} \rangle \rightarrow \text{if} (\langle \text{Expr} \rangle) \langle \text{Stmt} \rangle$
$\langle \text{StmtList} \rangle \rightarrow \langle \text{Stmt} \rangle$
$\langle \text{StmtList} \rangle \rightarrow \langle \text{StmtList} \rangle \langle \text{Stmt} \rangle$
$\langle \text{Expr} \rangle \rightarrow \langle \text{Id} \rangle$
$\langle \text{Expr} \rangle \rightarrow \langle \text{Num} \rangle$
$\langle \text{Expr} \rangle \rightarrow \langle \text{Expr} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle$
$\langle \text{Id} \rangle \rightarrow x$
$\langle \text{Id} \rangle \rightarrow y$
$\langle \text{Num} \rangle \rightarrow 0$
$\langle \text{Num} \rangle \rightarrow 1$
$\langle \text{Num} \rangle \rightarrow 9$
$\langle \text{Optr} \rangle \rightarrow >$
$\langle \text{Optr} \rangle \rightarrow +$

$\langle \text{Stmt} \rangle$		
if ($\langle \text{Expr} \rangle$	$\langle \text{Stmt} \rangle$
if ($\langle \text{Expr} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle$	$\langle \text{Stmt} \rangle$
if ($\langle \text{Id} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle$	$\langle \text{Stmt} \rangle$
if ($x \langle \text{Optr} \rangle \langle \text{Expr} \rangle$	$\langle \text{Stmt} \rangle$
if ($x > \langle \text{Expr} \rangle$	$\langle \text{Stmt} \rangle$
if ($x > \langle \text{Num} \rangle$	$\langle \text{Stmt} \rangle$
if ($x > 9$	$\langle \text{Stmt} \rangle$
if ($x > 9 \{ \langle \text{StmtList} \rangle \}$	$\langle \text{StmtList} \rangle$
if ($x > 9 \{ \langle \text{StmtList} \rangle \langle \text{Stmt} \rangle$	$\langle \text{Stmt} \rangle$
if ($x > 9 \{ \langle \text{Id} \rangle = \langle \text{Expr} \rangle ;$	$\langle \text{Stmt} \rangle$
if ($x > 9 \{ x = \langle \text{Expr} \rangle ;$	$\langle \text{Stmt} \rangle$
if ($x > 9 \{ x = \langle \text{Num} \rangle ;$	$\langle \text{Stmt} \rangle$
if ($x > 9 \{ x = 0 ;$	$\langle \text{Stmt} \rangle$
if ($x > 9 \{ x = 0 ; \langle \text{Id} \rangle = \langle \text{Expr} \rangle ;$	$\langle \text{Expr} \rangle$
if ($x > 9 \{ x = 0 ; y = \langle \text{Expr} \rangle$	$\langle \text{Expr} \rangle$
if ($x > 9 \{ x = 0 ; y = \langle \text{Expr} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle$	$\langle \text{Expr} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle$
if ($x > 9 \{ x = 0 ; y = \langle \text{Id} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle$	$\langle \text{Id} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle$
if ($x > 9 \{ x = 0 ; y = y \langle \text{Optr} \rangle \langle \text{Expr} \rangle$	$y \langle \text{Optr} \rangle \langle \text{Expr} \rangle$
if ($x > 9 \{ x = 0 ; y = y + \langle \text{Expr} \rangle$	$y + \langle \text{Expr} \rangle$
if ($x > 9 \{ x = 0 ; y = y + \langle \text{Num} \rangle$	$y + \langle \text{Num} \rangle$
if ($x > 9 \{ x = 0 ; y = y + 1 ;$	$y + 1 ;$

语法分析阶段的主题之一: 上下文无关文法

```

<Stmt> → <Id> = <Expr> ;
<Stmt> → { <StmtList> }
<Stmt> → if ( <Expr> ) <Stmt>
<StmtList> → <Stmt>
<StmtList> → <StmtList> <Stmt>
<Expr> → <Id>
<Expr> → <Num>
<Expr> → <Expr> <Optr> <Expr>
  <Id> → x
  <Id> → y
  <Num> → 0
  <Num> → 1
  <Num> → 9
  <Optr> → >
  <Optr> → +

```

语法分析阶段的主题之二: 构建语法分析树

$\langle \text{Stmt} \rangle$			
if ($\langle \text{Expr} \rangle$)	$\langle \text{Stmt} \rangle$
if ($\langle \text{Expr} \rangle$ $\langle \text{Optr} \rangle$ $\langle \text{Expr} \rangle$)	$\langle \text{Stmt} \rangle$
if ($\langle \text{Id} \rangle$ $\langle \text{Optr} \rangle$ $\langle \text{Expr} \rangle$)	$\langle \text{Stmt} \rangle$
if (x $\langle \text{Optr} \rangle$ $\langle \text{Expr} \rangle$)	$\langle \text{Stmt} \rangle$
if (x > $\langle \text{Expr} \rangle$)	$\langle \text{Stmt} \rangle$
if (x > $\langle \text{Num} \rangle$)	$\langle \text{Stmt} \rangle$
if (x > 9)	$\langle \text{Stmt} \rangle$
if (x > 9	{	$\langle \text{StmtList} \rangle$ }
if (x > 9	{	$\langle \text{StmtList} \rangle$ $\langle \text{Stmt} \rangle$ }
if (x > 9	{	$\langle \text{Stmt} \rangle$ $\langle \text{Stmt} \rangle$ }
if (x > 9	{	$\langle \text{Id} \rangle = \langle \text{Expr} \rangle ;$ $\langle \text{Stmt} \rangle$ }
if (x > 9	{	x = $\langle \text{Expr} \rangle ;$ $\langle \text{Stmt} \rangle$ }
if (x > 9	{	x = $\langle \text{Num} \rangle ;$ $\langle \text{Stmt} \rangle$ }
if (x > 9	{	x = 0 $\langle \text{Stmt} \rangle$ }
if (x > 9	{	x = 0 ; $\langle \text{Id} \rangle = \langle \text{Expr} \rangle ;$ }
if (x > 9	{	x = 0 ; y = $\langle \text{Expr} \rangle ;$ }
if (x > 9	{	x = 0 ; y = $\langle \text{Expr} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle ;$ }
if (x > 9	{	x = 0 ; y = $\langle \text{Id} \rangle \langle \text{Optr} \rangle \langle \text{Expr} \rangle ;$ }
if (x > 9	{	x = 0 ; y = y $\langle \text{Optr} \rangle \langle \text{Expr} \rangle ;$ }
if (x > 9	{	x = 0 ; y = y + $\langle \text{Expr} \rangle ;$ }
if (x > 9	{	x = 0 ; y = y + $\langle \text{Num} \rangle ;$ }
if (x > 9	{	x = 0 ; y = y + 1 ; }

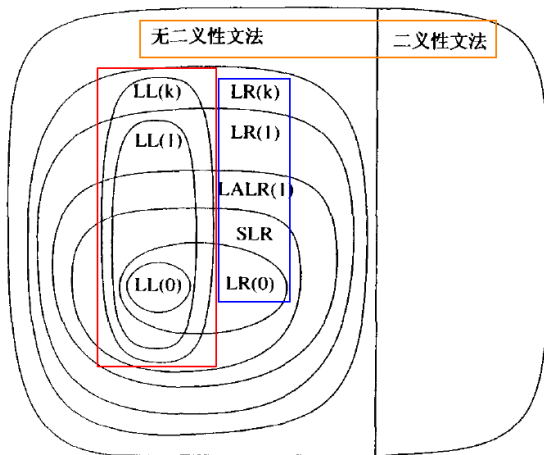
语法分析阶段的主题之三: 错误恢复



报错、**恢复**、继续分析

只考虑**无二义性**的文法

这意味着, 每个句子对应唯一的一棵语法分析树



今日份主题: **$LR(1)$ ($LR(0)$) 语法分析器**

自顶向下的、
不断规约的、
基于句柄查找自动机的、
适用于 $LR(1)$ ($LR(0)$) 文法的、
 $LR(1)$ ($LR(0)$) 语法分析器

自底向上构建语法分析树

根节点是文法的起始符号 S

叶节点是词法单元流 $w\$$

仅包含终结符号与特殊的文件结束符 $\$$

自底向上构建语法分析树

根节点是文法的起始符号 S

每个中间非终结符节点表示使用它的某条产生式进行归约

叶节点是词法单元流 $w\$$

仅包含终结符号与特殊的文件结束符 $\$$

“推导”与“归约”

从产生式的角度看, 是“推导”

$$A \rightarrow \alpha$$

从输入的角度看, 是“归约”

$$A \leftarrow \alpha$$

$$S \triangleq \gamma_0 \Rightarrow \dots \gamma_{i-1} \Rightarrow \gamma_i \Rightarrow \gamma_{r+1} \Rightarrow \dots \Rightarrow r_n = w$$

$$S \triangleq \gamma_0 \Leftarrow \dots \gamma_{i-1} \Leftarrow \gamma_i \Leftarrow \gamma_{r+1} \Leftarrow \dots \Leftarrow r_n = w$$

自底向上语法分析器为输入构造**反向推导**

$LR(*)$ 语法分析器

L : 从左向右 (left-to-right) 扫描输入

R : 构建反向 (reverse) 最右 (leftmost) 推导
在最右推导中, 最左叶节点最后才被处理

在反向最右推导中, 最左叶节点最先被处理 (与从左到右扫描一致)

(Q : 何时移入, 何时归约)

$$S' \rightarrow S\$$$

$$S \rightarrow (L)$$

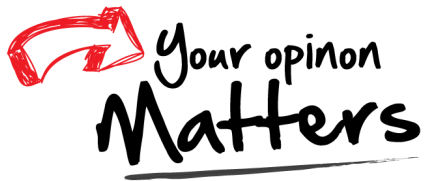
$$S \rightarrow x$$

$$L \rightarrow S$$

$$L \rightarrow L, S$$

$$(x, x, x)$$

Thank
You!



Office 926

hfwei@nju.edu.cn