

# 《编译原理》 实验报告

 实验名称:
 专题五基于 SLR(1)语义分析中间代码生成

 学
 号:

 22281188

 姓
 名:

 工家玮

 学
 院:

 计算机科学与技术学院

 日
 期:

 2024年12月17日

# 景目

一、 程序功能描述	1
1.1 理论知识	1
1.2 实验项目	1
1.3 设计说明	1
1.4 设计要求	1
二、主要数据结构描述	2
三、程序结构描述	2
3.1 Fisrt 集的计算	2
3.2 FOLLOW 集的计算	3
3.3 SLR 分析表的构造	4
3.4 四元式的生成	5
3.5 程序流程图如下:	6
3.6 函数定义以及函数之间的调用关系	6
四、程序测试	7
4.1 用例 1 测试	7
4.2 用例 2 测试	9
五、研究性教学实验感受	10
六、附件(源代码列表)	11

# 一、 程序功能描述

## 1.1 理论知识

语法制导的基本概念,目标代码结构分析的基本方法,赋值语句语法制导生 成四元式的基本原理和方法,该过程包括语法分析和语义分析过程。

## 1.2 实验项目

完成以下描述赋值语句 SLR(1)文法语法制导生成中间代码四元式的过程。

 $G[S]:S \rightarrow V = E$ 

 $E \rightarrow E + T | E - T | T$ 

 $T \rightarrow T*F|T/F|F$ 

 $F \rightarrow (E)|i$ 

V→i

## 1.3 设计说明

终结符号i为用户定义的简单变量,即标识符的定义。

## 1.4 设计要求

- (1) 构造文法的 SLR(1)分析表,设计语法制导翻译过程,给出每一产生式对应的语义动作;
- (2) 设计中间代码四元式的结构;
- (3)输入串应是词法分析的输出二元式序列,即某赋值语句"专题 1"的输出结果,输出为赋值语句的四元式序列中间文件;
- (4)设计两个测试用例(尽可能完备),并给出程序执行结果四元式序列。
- (5) 考虑根据文法自动构造 SLR(1)分析表,并添加到你的程序中。

# 二、主要数据结构描述

变量类型	变量名称	说明
string	start	 开始符号
set <string></string>	Vt	终结符号集
set <string></string>	Vn	非终结符号集
set <string></string>	VtP	终结符号+#
map <string, set<string="">&gt;</string,>	first	first 集
map <string, set<string="">&gt;</string,>	follow	follow 集
struct	struct status	DFA 活前缀
struct	ActionNode	Action 结点
struct	GotoNode	Goto 结点
list <actionnode></actionnode>	Action	Action
list <gotonode></gotonode>	Goto	Goto
vector <ananode></ananode>	stack	分析栈
vector <string></string>	leftString	剩余输入串
vector <string></string>	temp	中转记录输入串
vector <string></string>	recValue	中转记录真实输入串
vector <string></string>	recleftValue	剩余真实输入串

# 三、程序结构描述

## **3.1 Fisrt 集的计算**

先遍历所有产生式,识别出右侧以终结符开头的产生式,例如 $A \to \alpha$ 。 在此过程中,必须特别注意不遗漏那些右侧仅由单一终结符组成的产生式,如 $A \to \alpha$ 。需要强调的是,许多研究者在遇到  $A \to \alpha$ 时,常误以为终结符 a 后必然跟随其他符号,这种假设是不准确的。因此,对于每一个符合条件的产生式,应将终结符 a 添加到 First(A) 集合中。 其次,继续遍历所有产生式,识别出右侧以非终结符开头的产生式,例如  $A \to B\alpha$ 。在此步骤中,同样需要注意处理那些右侧仅包含非终结符 B的产生式,如 $A \to B$ 。对于每一个此类产生式,应将 FIRST(B) 集合中的所有终结符添加到 FIRST(A) 集合中。

## 3.2 FOLLOW 集的计算

- 1. 将终结符 # 加入到 FOLLOW(S) 中, 其中 S 为文法的起始符号。
- 2. 遍历所有产生式,识别右侧包含子串 Ba... 的产生式,例如  $A \rightarrow \alpha B \alpha \beta$ 。在此过程中,需特别注意 B 前后可能不存在任何符号的情况。对于每一个符合条件的产生式,应将终结  $\alpha$  添加到FOLLOW(B)中。
- 3. 遍历所有产生式,识别右侧包含相邻非终结符对 BC 的产生式,例如  $A \to \alpha BCDE$  。对于任意两个相邻的非终结符(B, C)、(C, D)、(D, E),应将 FIRST(C)、FIRST(D)、FIRST(E) 中除空串以外的所有终结符加入到 FOLLOW(B)、FOLLOW(C)、FOLLOW(D) 中。
- 4. 遍历所有产生式,识别右侧以单一非终结符 B 结尾的产生式,例如  $A \to \alpha B$ 。此时,应将FOLLOW(A) 集合中的所有终结符添加到 FOLLOW(B) 中。特别需要注意的是,不应遗漏那些右侧仅包含一个非终结符的产生式,如 $A \to B$ 。
- 5. 遍历所有产生式,识别右侧以相邻非终结符对 BC 结尾的产生式,例如  $A \to \alpha BCDE$ 。在此情形下,需考虑最后两个非终结符 D 和 E 的关系。若 FIRST(E) 包含空串,则应将FOLLOW(A)添加到 FOLLOW(D) 中。此步骤实 质上是步骤 4 的一个特殊情况。

#### (3) 有效项目集规范族的构造

- 1. 闭包函数*CLOSURE(I)*的定义如下:
- 初始条件: 所有属于集合 I 的项目均包含在 CLOSURE(I) 中。
- 扩展条件: 若  $A \to \alpha \cdot B\beta$  属于CLOSURE(I),则每一个形如  $B \to \cdot \gamma$  的项目也必须包含在 CLOSURE(I) 中。
- 收敛条件: 重复应用扩展条件,直至不再有新的项目被加入,即 CLOSURE(I)不再扩展。

2. 转换函数 GO(I, X) 的定义:

$$GO(I,X) = CLOSURE(J)$$

其中:

- / 为包含某一项目的状态集合。
- X为文法符号,且 $X \in (V_N \cup V_T)$ 。
- $J = \{ A \rightarrow \alpha X \cdot \beta | A \rightarrow \alpha \cdot X \beta \in I \}$

通过应用闭包函数和转换函数,可以构造文法  $G\{'\}$  的 LR(0) 项目集规范族,具体步骤如下:

- a) 初始状态的构造:将项目  $S' \to \cdot S$  作为初态集的核心,然后对其求闭包,得到初态的项目集 $CLOSURE(\{S' \to \cdot S\})$ 。
- b) 状态的扩展:对初态集及后续构造的各个项目集应用转换函数 GO(I,X) = CLOSURE(J),以求得新的状态集J。
- c) 迭代过程: 重复步骤 b, 直至所有可能的状态均被构造, 且不再生成新的状态。

## 3.3 SLR 分析表的构造

首先,将文法 G 拓广为G',并基于G'构造 LR(0) 项目集规范族 C 及其状态转换函数GO。随后,依据以下方法构造分析表中的ACTION和 GOTO 函数:

#### 1. 移进操作:

✓ 若项目 $A \to \alpha \cdot b$  β属于 $I_k$ ,且  $GO(I_k, b) = I_j$  (其中 b 为终结符),则将ACTION[k, b] 设为"移进至状态 j",记为" $s_j$ "。

#### 2. 规约操作:

✓ 若项目  $A \to \alpha$  · 属于 $I_k$  ,则对所有  $a \in FOLLOW(A)$  ,将 ACTION[k, a] 设为"使用产生式  $A \to \alpha$  进行规约",记为"  $r_j$ ",其中假定  $A \to \alpha$  为 文法 G 的第 j 个产生式。

#### 3. 接受操作:

✓ 若项目 $S' \to S$ · 属于 $I_k$ , 则将 ACTION $[k, \ \#]$ 设为"接受",记为"acc"。

#### 4. GOTO 操作:

- ✓ 若  $GO(I_k, A) = I_i$  (其中 A 为非终结符),则将 GOTO[k, A] 设为 j 。
- ✓ 在构造分析表的过程中,凡是未被上述规则覆盖的空白格应标记为"出错标志"。
- ✓ 分析器的初始状态为包含  $S' \rightarrow S'$  的项目集状态。
- ✓ SLR 分析中可能遇到的冲突类型主要包括移进-规约冲突和规约-规约冲突。

#### 5. 冲突检测

✓ 通过 Conflict 函数检测当前状态的 FIRST 集合和 FOLLOW 集合是否存在 冲突。如果检测到冲突,应采取相应措施予以消除;若无冲突,则继续 执行后续步骤。

## 3.4 四元式的生成

根据语法分析过程进行语法制导翻译,具体步骤如下:

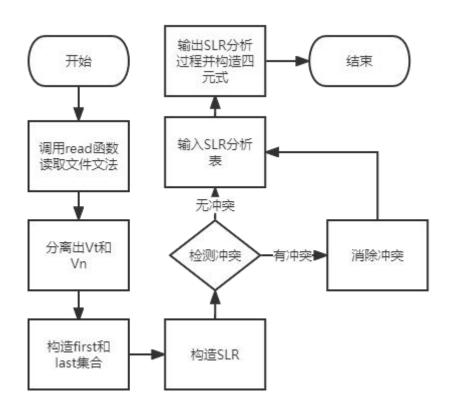
- 1. 栈的设立:设立三个栈,分别为状态栈、符号栈以及存放四元式变量的栈。
  - 2. 规约过程:
  - 当出现规约操作  $V \rightarrow \alpha$  时,将  $\alpha$  压入相应的栈中。
  - 对于特定的产生式规则,例如:
  - $A \rightarrow V = E$
  - $E \rightarrow E + T$  或  $E \rightarrow E T$
  - $T \rightarrow T * F \otimes T / F$

使用 newTemp 函数生成新的中间变量,根据语义规则生成相应的四元式。

- 3. 变量的管理:
- 在规约过程中,将对应栈中的变量弹出,并将新生成的变量压入四元式变量栈中。

通过上述步骤,可以生成反映程序中间代码逻辑的四元式,从而为后续的代码生成和优化提供基础。

## 3.5 程序流程图如下:



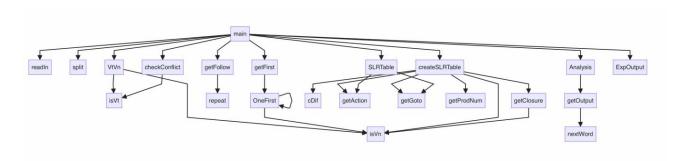
# 3.6 函数定义以及函数之间的调用关系

## 3.6.1 函数定义

函数名称	函数功能描述
split()	分解文法规则,将每条产生式拆分为左部和
	右部,并根据
isNumber()	判断字符串是否为数字,返回布尔值。
readIn()	计算所有非终结符的 FIRSTVT 集合,用于
	构建优先关系表,确定终结符之间的优先
	级。
VtVn()	根据读取的文法规则区分终结符和非终结
	符,并将 # 加入扩展终结符集合。
OneFirst()	递归计算单个非终结符的 FIRST 集,处理
	终结符和非终结符的 FIRST 规则,并存储
	结果。
getFirst()	计算所有非终结符的 FIRST 集,并输出结
	果。
getFollow()	计算所有非终结符的 FOLLOW 集,按照规
	则逐步更新,并输出结果。

cDif()	比较两个状态是否相等,用于判定 DFA 状	
	态是否重复。	
getClosure()	计算一个状态的闭包, 处理项集中的	
	点 · 后的符号, 递归生成新状态项。	
createSLRTable()	创建 SLR(1) 分析表,包含 ACTION 和	
	GOTO 表的生成,同时输出项集族。	

## 3.6.2 函数之间的调用关系



# 四、程序测试

## 4.1 用例 1 测试

## 测试用例:

(identifier,p)
(=, )
(identifier,x)
(+,)
((,)
(identifier,x)
(/,)
((,)
(identifier,y)
(+,)
(identifier,z)
(0,)
((,)
(*,)
((,)
(identifier,z)
(/,)
(identifier,y)
(*,)

```
(identifier,y)
(),)
```

(1) 生成的 First 和 Follow 集如下

```
First:
     A:i
     E:(
                         i
     F:(
           i
     S:i
     T:(
                  i
           *
     V:i
Follow:
     A:#
     E:#
     T:#
     V:=
```

(2) 生成的有效项目集规范族如下

```
getClosure:

C0={ S->.A, V->.i, A->.V=E }

C1={ S->A. }

C2={ V->i. }

C3={ A->V.=E }

C4={ A->V=.E, E->.E+T, E->.E-T, E->.T, T->.T*F, T->.T/F, F->.(E), F->.i }

C5={ A->V=E., E->E.+T, E->E.-T }

C6={ E->T., T->T.*F, T->T./F }

C7={ T->F. }

C8={ F->(.E), E->.E+T, E->.E-T, E->.T, T->.T*F, T->.T/F, F->.(E), F->.i }

C9={ F->i. }

C10={ E->E-.T, F->.(E), F->.i, T->.F, T->.T*F, T->.T/F }

C11={ E->E-.T, F->.(E), F->.i }

C12={ T->T*.F, F->.(E), F->.i }

C13={ T->T/.F, F->.(E), F->.i }

C14={ F->(E.), E->E+T, E->E.-T }

C15={ E->E+T., T->T.*F, T->T./F }

C15={ E->E-T., T->T.*F, T->T./F }

C17={ T->T*F. }

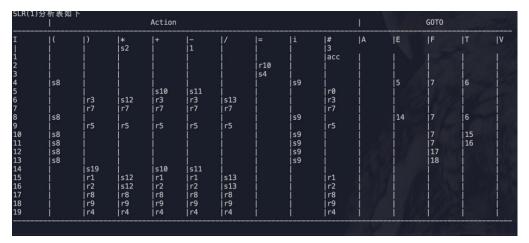
C18={ T->T/F. }

C19={ F->(E). }
```

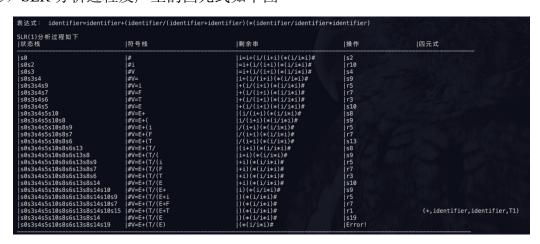
(3) 冲突检测结果如下:

含有冲突: I5有冲突: I6有冲突: I15有冲突: I16有冲突:

(4) SLR(1)分析表构造如下



(5) SLR 分析过程及产生的四元式如下图



(6) 最终将中间代码输入到文件中如下图

```
(+,identifier,identifier,T1)
```

## 4.2 用例 2 测试

#### 测试用例:

```
(identifier,p)
(=,)
(identifier,x)
(+,)
((,)
(identifier,x)
(/,)
((,)
(identifier,y)
(+,)
```

```
(identifier,z)
(),)
(),)
(*,)
((,)
(identifier,z)
(/,)
(identifier,y)
(*,)
(identifier,y)
(),)
```

First 集、Follow 集、SLR 分析表的构造都同 5.1。 测试结果如下图,

R(1)分析过程如下 术态栈	符号栈	剩余串	操作	四元式
 0	#	i=i+(i/(i+i))*(i/i*i)#	s2	PERSONAL PROPERTY OF THE PROPE
0s2	#i	=i+(i/(i+i))*(i/i*i)#	r10	
0s3	1#V	=i+(i/(i+i))*(i/i*i)#	54	
0s3s4	#V=	i+(i/(i+i))*(i/i*i)#	ls9	
0s3s4s9	I#V=i	+(i/(i+i))*(i/i*i)#	r5	
0s3s4s7	I#V=F	+(i/(i+i))*(i/i*i)#	ir7	
0s3s4s6	I#V=T	+(i/(i+i))*(i/i*i)#	ir3	
)s3s4s5	I#V=E	+(i/(i+i))*(i/i*i)#	s10	
s3s4s5s10	#V=E+	(i/(i+i))*(i/i*i)#	is8	
s3s4s5s10s8	#V=E+(	i/(i+i))*(i/i*i)#	is9	
s3s4s5s10s8s9	#V=E+(i	/(i+i))*(i/i*i)#	ir5	
)s3s4s5s10s8s7	#V=E+(F	//(i+i))*(i/i*i)#	r7	
)s3s4s5s10s8s6	#V=E+(T	//(i+i))*(i/i*i)#	Is13	
s3s4s5s10s8s6s13	#V=E+(T/	(i+i))*(i/i*i)#	Is8	
9s3s4s5s10s8s6s13s8	#V=E+(T/(	i+i))*(i/i*i)#	is9	
)s3s4s5s10s6s6s13s8s9	#V=E+(T/(i	+i))*(i/i*i)#	lr5	
)s3s4s5s10s0s0s13s0s <i>3</i>	#V=E+(T/(F	+i/)*(i/i*i)#	r7	
)s3s4s5s10s6s6s13s8s6	#V=E+(T/(T	+1/)*(1/1*1)#  +1))*(1/1*1)#	lr3	
)s3s4s5s10s6s6s13s6s6	#V=E+(T/(E	+1/)*(1/1*1)#  +1))*(1/1*1)#	s10	
)s3s4s5s10s6s6s13s6s14 )s3s4s5s10s8s6s13s8s14s10	#V=E+(T/(E+	i))*(i/i*i)#	s10  s9	
0535455510505051350514510 053545551058565135851451059	#V=E+(T/(E+i	1	r5	
053545351058565135851451059 053545551058565135851451057	#V=E+(T/(E+F	))*(1/1*1)#  ))*(1/1*i)#	r7	
053545351056565135651451057 0535455510585651358514510515	#V=E+(T/(E+T	))*(1/1*1)#  ))*(1/1*i)#	ri	(+,identifier,identifier,T1
s3s4s5s10s8s6s13s8s14s10s15	#V=E+(T/(E+1	))*(1/1*1)#  ))*(1/1*i)#	s19	(+,identifier,identifier,ii
0535455510565651356514 0535455510585651358514519			S19   r4	
9535453510565651356514519 95354555105856513518	#V=E+(T/(E)  #V=E+(T/F	)*(i/i*i)#  )*(i/i*i)#	119	(/ identifier T1 T2)
95354535105656513516 95354555105856	#V=E+(T/F	)*(1/1*1)#  )*(1/1*1)#	Ir3	(/,identifier,T1,T2)
05354535105656 053545551058514	#V=E+(E	)*(1/1*1)#  )*(1/1*1)#	s19	
953545351050514 953545551058514519	#V=E+(E)	/*(1/1*1)#  *(1/1*1)#	S19   r4	
053545351050514519 053545551057	#V=E+(E)  #V=E+F		r7	
		*(i/i*i)#		
ls3s4s5s10s15	#V=E+T	*(i/i*i)#	s12	
)s3s4s5s10s15s12	#V=E+T*  #V=E+T*(	(i/i*i)#  i/i*i)#	s8  s9	
s3s4s5s10s15s12s8				
s3s4s5s10s15s12s8s9	#V=E+T*(i	/i*i)#	ir5	
s3s4s5s10s15s12s8s7	#V=E+T*(F	/i*i)#	r7	
s3s4s5s10s15s12s8s6	#V=E+T*(T	/i*i)#	s13	
s3s4s5s10s15s12s8s6s13	#V=E+T*(T/	i*i)#	s9	
s3s4s5s10s15s12s8s6s13s9	#V=E+T*(T/i	*i)#	r5	The state of the s
s3s4s5s10s15s12s8s6s13s18	#V=E+T*(T/F	*i)#	r9	(/,identifier,identifier,T3
s3s4s5s10s15s12s8s6	#V=E+T*(T	*i)#	s12	
s3s4s5s10s15s12s8s6s12	#V=E+T*(T*	i)#	s9	
s3s4s5s10s15s12s8s6s12s9	#V=E+T*(T*i	1)#	r5	The state of the s
s3s4s5s10s15s12s8s6s12s17	#V=E+T*(T*F	[)#	r8	(*,T3,identifier,T4)
s3s4s5s10s15s12s8s6	#V=E+T*(T	1)#	r3	
s3s4s5s10s15s12s8s14	#V=E+T*(E	])#	s19	
s3s4s5s10s15s12s8s14s19	#V=E+T*(E)	#	r4	
s3s4s5s10s15s12s17	#V=E+T*F	<b> </b> #	r8	(*,T2,T4,T5)
)s3s4s5s10s15	#V=E+T	#	r1	(+,identifier,T5,T6)
ðs3s4s5	#V=E	į#	r0	(=,T6,NULL,identifier)
)s1	I#A	i#	lacc	

# 五、研究性教学实验感受

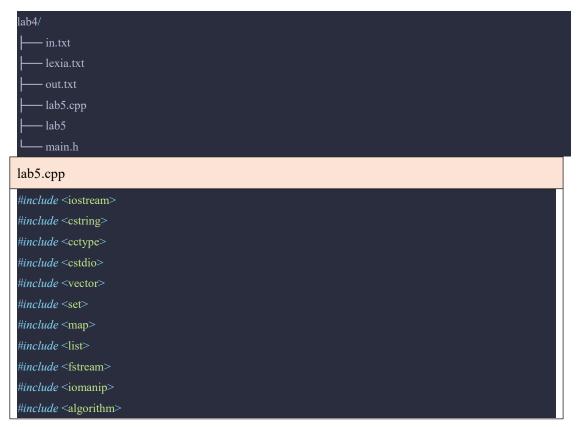
通过完成《编译原理》中"SLR(1)分析法的语法制导翻译及中间代码生成"的实验,我深刻体会到理论与实践结合的重要性。这次实验涵盖了从文法的FIRST和FOLLOW集计算到SLR分析表的构造,再到中间代码的生成整个过程,让我对编译器设计的核心环节有了更直观的认识。从理论层面上,我不仅理解了

SLR(1)文法的分析方法,还进一步掌握了语法制导翻译的原理以及四元式中间代码的生成方式。在实践过程中,通过动手设计数据结构并实现算法逻辑,我真正体会到了编译原理从抽象到具体的转化过程。

实验的完成让我学会了如何将复杂问题逐步分解并实现。从初始的分析表设计,到冲突检测与优化,再到语义动作的设置和四元式生成,我逐步完善了实验的各个环节。在此过程中,我不仅增强了编程能力,还培养了面对问题时的逻辑推理与创新思维。此外,测试用例的设计和运行结果的分析也让我意识到程序健壮性和完备性的关键作用。在不断调试、优化代码的过程中,我更加体会到持续改进的重要性

实验还带给我一种从理论实现到实际结果的成就感。当程序运行并生成正确的四元式中间代码时,我切实感受到自己对编译原理的理解在逐步深化。整个实验不仅让我在技术上有所提升,也让我更深刻地认识到编译器设计中的严谨性和创造性,这为我后续的学习和研究奠定了坚实的基础。

# 六、附件(源代码列表)



```
#include<string>
#include<iterator>
#include <sstream>
#pragma warning(disable:4996)
using namespace std;
set<string> prod; // 产生式集合
map<string, set<string> > splProd; // 分解的产生式集合
string start; // 开始符号
set<string> Vt; // 终结符集合
set<string> Vn; // 非终结符集合
set<string> VtPlus; // 终结符集合 +#
map<string, set<string> > first; // first 集
map<string, set<string> > follow; // follow 集
string ptr; // 当前词法单元对应的字符
string RecValue; // 当前词法单元对应的实际值
struct Item {
struct status {
    list<Item> Itemlist;
}; // 状态
list<status> DFA;
Item firstItem; // 初始项目10
struct ActionNode {
struct GotoNode {
list<ActionNode> ACTION;
list<GotoNode> GOTO;
struct varX {
```

```
string X; // 符号
string value; // 符号的值/Temp
}; // 符号
struct anaNode {
    int S; // 状态
    varX X; // 符号
};
vector<anaNode> stack; // 分析稅(符号稅和状态稅)
vector<string> leftString; // 剩余输入串
vector<string> temp; // 临时记录输入串
vector<string> recValue; // 絕時记录彙际输入串
vector<string> recleftValue; // 剩余类际输入串
struct quaterExp {
    string arg1;
    string arg2;
    string arg2;
    string result;
};
list<quaterExp> Exp; // 四元式
list<int> quaterI; // 可产生四元式的状态
int recTemp = 0; // 记录 result 编号
```

#### main.h

```
#include <iostream>
#include <cstring>
#include <cctype>
#include <cstdio>
#include <vector>
#include <set>
#include <map>
#include <list>
#include <fstream>
#include <iomanip>
#include <algorithm>
#include<string>
#include<iterator>
#include <sstream>
#pragma warning(disable:4996)
using namespace std;
set<string> prod; // 产生式集合
map<string, set<string>> splProd; // 分解的产生式集合
```

```
string start; // 开始符号
set<string> Vt; // 终结符集合
set<string> Vn; // 非终结符集合
set<string> VtPlus; // 终结符集合 +#
map<string, set<string>> first; // first 集
map<string, set<string> > follow; // follow 集
string ptr; // 当前词法单元对应的字符
string RecValue; // 当前词法单元对应的实际值
struct Item {
struct status {
    list<Item> Itemlist;
}; // 状态
list<status> DFA;
Item firstItem; // 初始项目10
struct ActionNode {
struct GotoNode {
list<ActionNode> ACTION;
list<GotoNode> GOTO;
struct varX {
}; // 符号
struct anaNode {
```

```
vector<anaNode> stack; // 分析核(符号核和状态核)
vector<atring> leftString; // 剩余输入串
vector<atring> recValue; // 临时记录输入串
vector<atring> recleftValue; // 剩余实际输入串
struct quaterExp {
    string op;
    string arg1;
    string arg2;
    string result;
};
list<quaterExp> Exp; // 四元式
list<int> quaterI; // 可产生四元式的状态
int recTemp = 0; // 记录 result 编号
```