

清华大学本科生考试试题专用纸

考试课程 《编译原理》 (A 卷) 2018 年 1 月 9 日

学号: _____ 姓名: _____ 班级: _____

(注: 解答可以写在答题纸上, 也可以写在试卷上; 交卷时二者均需上交。)

一. (12分) 必做实验相关简答题 (所得分数直接加到总评成绩)

二. (12分)

以下是某简单语言的一段代码。语言中不包含数据类型的声明, 所有变量的类型默认为整型 (假设占用一个存储单元)。语句块的括号为 'begin' 和 'end' 组合; 赋值号为 ':='。每一个过程声明对应一个静态作用域。该语言支持嵌套的过程声明, 但只能定义无参过程, 且没有返回值。过程活动记录中的控制信息包括静态链 SL, 动态链 DL, 以及返回地址 RA。程序的执行遵循静态作用域规则。

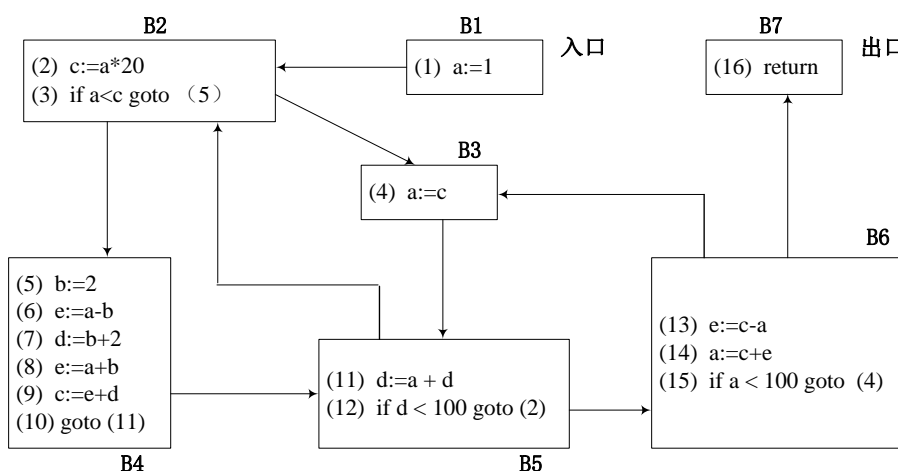
```
(1)  var a0,b0;
(2)  procedure fun1 ;
(3)      var a1,b1;
(4)      procedure fun2 ;
(5)          var a2;
(6)          begin
(7)              a2:= a1*a0-b1;
(8)              if(a2<0) then call fun3;
.              ..... /*不含任何 call 语句和声明语句*/
.          end;
.      begin
.          a1:= a0 - b0;
.          b1:=a0 + b0;
(x)      If  a1 < a0  then  call fun2 ;
.          ..... /*不含任何 call 语句和声明语句*/
.      end ;
.  procedure fun3 ;
.      var a3,b3;
.      begin
.          a3:=a0*b0 ;
.          b3:=a0/b0 ;
(y)      if(a3 <> b3) call fun1 ;
.          ..... /*不含任何 call 语句和声明语句*/
.      end ;
.  begin
.      a0 := 1;
.      b0 := 2;
.      call fun3;
.      ..... /*不含任何 call 语句和声明语句*/
.  end .
```

1. (6分) 若实现该语言时符号表的组织采用多符号表结构, 即每个静态作用域均对应

一个符号表。在静态语义分析中，为了体现作用域信息，须要维护一个作用域栈，假设栈中的元素是指向某个作用域的指针。试指出：在分析至语句（7）时，当前作用域栈中包含的作用域的指针有多少个？分别是哪几个作用域？其中，次栈顶所指向的作用域中包含哪些符号？

2. （6分）当过程fun1被第二次激活时，运行栈上共有几个活动记录？分别是那些过程的活动记录？当前位于栈顶和次栈顶的活动记录中静态链 SL 和动态链 DL 分别指向什么位置？（注：指出是哪个活动记录的起始位置即可）

三 (13分) 下图是包含 7 个基本块的流图，其中 B1 为入口基本块，B7 为出口基本块：



1. (3分) 指出在该流图中存在的回边，以及该回边所对应的自然循环（即指出循环中所包含的基本块）。

2. (4分) 已知基本块 B2 和 B6 入口处的活跃变量（live variables）信息分别为

$$\text{LiveIn}(B2) = \{a, d\} \quad \text{和} \quad \text{LiveIn}(B6) = \{a, c, d\}$$

试计算 $\text{LiveIn}(B5) = ?$ 并指出在基本块 B4 内第 (7) 条语句之前处的活跃变量信息。

3. (3分) 已知基本块 B2 出口处的到达-定值（reaching definitions）信息为

$$\text{Out}(B2) = \{1, 2, 4, 5, 8, 11, 13\}$$

试指出在基本块 B4 内第 (8) 条语句使用变量 a 的 UD 链。

4. (3分) 试从基本块 B4 的 DAG 图，导出一个算术表达式，用来表示结点 c 的计算结果。要求该表达式中的运算数仅包含 DAG 图的叶子结点。（注：基本块入口处活跃变量所对应的叶子结点可表示为 a_0, b_0, c_0, \dots ）

四 (18 分) 给定某类表达式文法 $G[E]$:

$$\begin{aligned} E &\rightarrow +ER \mid -ER \mid \textit{positive} R \\ R &\rightarrow *ER \mid \varepsilon \end{aligned}$$

其中，+ 和 - 分别代表一元正和一元负运算，* 代表普通的二元乘法运算，positive 为代表正整数（非0）的单词。

1. (3分) 针对文法 $G[E]$ ，下表给出各产生式右部文法符号串的 $First$ 集合，各产生式左部非终结符的 $Follow$ 集合，以及各产生式的预测集合 PS 。试填充其中空白表项（共3处）的内容：

$G[E]$ 的规则 r	$First(rhs(r))$	$Follow(lhs(r))$	$PS(r)$
$E \rightarrow +ER$	+		+
$E \rightarrow -ER$	-	此处不填	-
$E \rightarrow \underline{positive} R$	<u>positive</u>	此处不填	<u>positive</u>
$R \rightarrow *ER$	*		*
$R \rightarrow \varepsilon$	ε	此处不填	

表中， $rhs(r)$ 为产生式 r 右部的文法符号串， $lhs(r)$ 为产生式 r 左部的非终结符。

2. (2分) $G[E]$ 不是 $LL(1)$ 文法，试解释为什么？
3. (6分) 虽然 $G[E]$ 不是 $LL(1)$ 文法，但可以采用一种强制措施，使得常规的 $LL(1)$ 分析算法仍然可用。针对含5个单词的输入串 $+ - 20 * 18$ ，以下基于这一措施以及上述各产生式的预测集合（或预测分析表）的一个表驱动 $LL(1)$ 分析过程：

步骤	下推栈	余留符号串	下一步动作
1	# E	$+ - 20 * 18 \#$	应用产生式 $E \rightarrow +ER$
2	# $RE +$	$+ - 20 * 18 \#$	匹配栈顶和当前输入符号
3	# RE	$- 20 * 18 \#$	应用产生式 $E \rightarrow -ER$
4	# $REE -$	$- 20 * 18 \#$	匹配栈顶和当前输入符号
5	# REE	$20 * 18 \#$	应用产生式 $E \rightarrow \underline{positive} R$
6	# $RRR \underline{positive}$	$20 * 18 \#$	匹配栈顶和当前输入符号
7	# RRR	$* 18 \#$	
8			
9			
10			
11			
12	# RRR	#	应用产生式 $R \rightarrow \varepsilon$
13	# RR	#	应用产生式 $R \rightarrow \varepsilon$
14	# R	#	应用产生式 $R \rightarrow \varepsilon$
15	#	#	结束

试填写上述分析过程中第7步时使用的产生式，以及第 8~11 步的分析过程，共计13处空白；并指出采用了什么样的强制措施。

4. (7分) 如下是以 $G[E]$ 为基础改造的一个 L 翻译模式：

$$\begin{aligned}
E &\rightarrow + E_1 \{ R.i := E_1.sign \} R \{ E.sign := R.s \} \\
E &\rightarrow - E_1 \{ R.i := \text{if } E_1.sign = P \text{ then } N \text{ else } P \} R \{ E.sign := R.s \} \\
E &\rightarrow \underline{positive} \{ R.i := P \} R \{ E.sign := R.s \} \\
R &\rightarrow * E \{ R_1.i := R.i \} R_1 \{ R.s := \text{if } E.sign = R_1.s \text{ then } P \text{ else } N \} \\
R &\rightarrow \varepsilon \{ R.s := R.i \}
\end{aligned}$$

其中，每个子表达式 E 的符号（即值大于零或小于零）记录在属性 $E.sign$ 中（属性值分别用 P 或 N 表示，其类型为枚举型 $SIGN = \text{enum} \{P, N\}$ ）。

试针对该 L 翻译模式构造一个自上而下的递归下降（预测）翻译程序（注：采用了上一小题中所述的强制措施）：

```

SIGN ParseE()                // 主函数
{
    switch (lookahead) {      // lookahead 为下一个输入符号
        case '+':
            MatchToken('+');
            e1_sign := ParseE();
            _____ ① _____ ;
            rs := ParseR(ri);
            e_sign := rs;
            break;
        case '-':
            MatchToken('-');
            e1_sign := ParseE();
            ri := if e1_sign=P then N else P;
            rs := ParseR(ri);
            e_sign := rs;
            break;
        case positive:
            MatchToken(positive);
            rs := ParseR(P);
            _____ ② _____ ;
            break;
        default:
            printf("syntax error \n")
            exit(0);
    }
    return e_sign;
}

SIGN ParseR(SIGN i)
{
    switch (lookahead) {      // lookahead 为下一个输入符号
        case _____ ③ _____:
            MatchToken('*');
            e_sign := ParseE();
            _____ ④ _____ ;
            r1s := ParseR(r1i);
            _____ ⑤ _____ ;
            break;
        case _____ ⑥ _____:
            _____ ⑦ _____ ;
            break;
    }
}

```

```

        default:
            printf("syntax error\n");
            exit(0);
    }
    return rs;
}

```

其中使用了与课程中所给的 MatchToken 函数。

该翻译程序中 ①~⑦ 的部分未给出，试填写之。

五 (25 分) 设语言 $L = \{ a^n b^m \mid m > n \geq 0 \}$, 下列 $G_1[S]$ 和 $G_2[S]$ 是该语言的两个无二义文法(对应两种不同的配对方法):

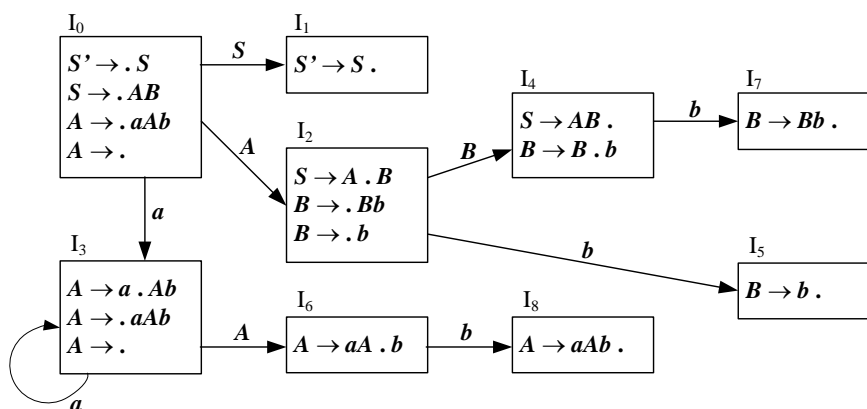
$G_1[S]$:

- (1) $S \rightarrow AB$
- (2) $A \rightarrow aAb$
- (3) $A \rightarrow \varepsilon$
- (4) $B \rightarrow Bb$
- (5) $B \rightarrow b$

$G_2[S]$:

- (1) $S \rightarrow aSb$
- (2) $S \rightarrow B$
- (3) $B \rightarrow Bb$
- (4) $B \rightarrow b$

1. (4分) 下图是相应于 $G_1[S]$ 的 LR(0) 自动机:

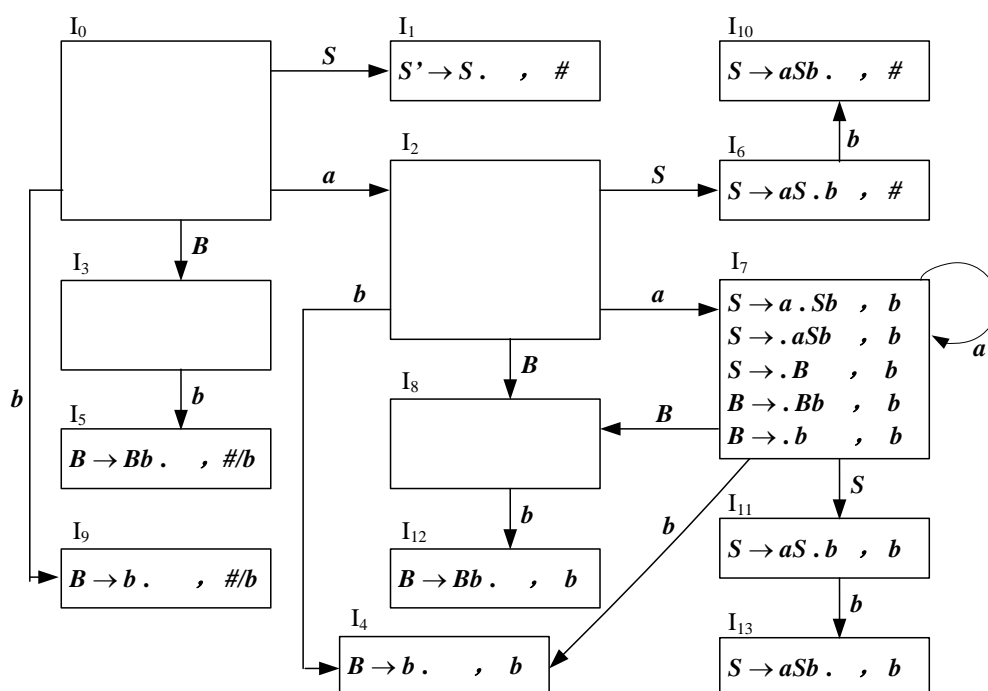


文法 $G_1[S]$ 不是 LR(0) 文法。试指出 $G_1[S]$ 的 LR(0) 自动机中存在哪些冲突的状态? 并指出这些状态的冲突类别, 即是移进-归约冲突还是归约-归约冲突?

3. (4分) 下图表示 $G_1[S]$ 的 LR(0) 分析表和 SLR(1) 分析表中状态 I_3 和 I_4 两行所对应的内容, 上半部分是 LR(0) 分析表, 下半部分是 SLR(1) 分析表, 但表中的 ACTION 部分没有给出, 试补齐之。

状态		ACTION			GOTO		
		<i>a</i>	<i>b</i>	#	<i>S</i>	<i>A</i>	<i>B</i>
LR(0) 分析表	3					6	
	4						
SLR(1) 分析表	3					6	
	4						

4. (4分) 在针对 $G_1[S]$ 的SLR(1)分析过程的某个时刻, 符号栈的栈顶是 A , 且栈中包含 a , 则此时所期待的句柄有哪些? 而在另一个时刻, 符号栈的栈顶是 A , 但栈中不包含 a , 则此刻所期待的句柄有哪些?
5. (2分) 在针对 $G_1[S]$ 的SLR(1)分析过程中, 若输入消耗掉相同数目的 a 和 b 时发生错误, 则最不可能的报错信息是 _____? 请在“末尾少 a ”, “末尾少 b ”, “末尾多 a ”和“末尾多 b ”中选择。
6. (4分) 下图是相应于 $G_2[S]$ 的LR(1)自动机, 但部分状态所对应的项目集未给出, 试补齐之 (即分别给出状态 I_0, I_2, I_3 和 I_8 对应的项目集)。



7. (3分) $G_2[S]$ 是否LR(1)文法? 若不是, 则指出 $G_2[S]$ 的LR(1)自动机中哪些状态有冲突?

六 (14 分)

给定如下文法 $G[S]$:

- (1) $S \rightarrow P$
- (2) $P \rightarrow P P \wedge$
- (3) $P \rightarrow P P \vee$
- (4) $P \rightarrow P \neg$
- (5) $P \rightarrow \underline{id}$

其中, \wedge 、 \vee 、 \neg 分别代表逻辑与、或、非等运算符单词, \underline{id} 代表标识符单词。如下是以 $G[S]$ 为基础文法的一个 L 翻译模式:

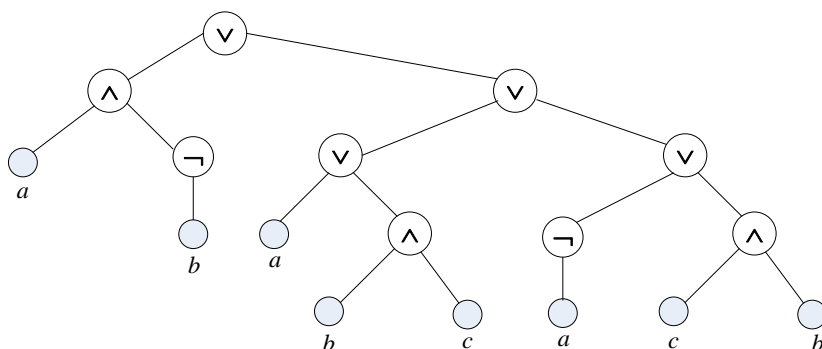
- (1) $S \rightarrow \{ P.i := 0 \} P \{ \text{print}(P.s) \}$
- (2) $P \rightarrow \{ P_1.i := P.i \} P_1 \{ P_2.i := P_1.s \} P_2 \wedge \{ P.s := P_2.s + 2 \}$
- (3) $P \rightarrow \{ P_1.i := P.i \} P_1 \{ P_2.i := P_1.s \} P_2 \vee \{ P.s := P_2.s + 1 \}$
- (4) $P \rightarrow \{ P_1.i := P.i \} P_1 \neg \{ P.s := P.i + P_1.s \}$
- (5) $P \rightarrow \underline{id} \{ P.s := 0 \}$

1. (3分) 对于合法输入串 $a b \neg \wedge a b c \wedge \vee a \neg c b \wedge \vee \vee \vee$, 该翻译模式的语义计算结果是什么? (回答 `print` 的执行结果即可, 这里 `print` 为普通显示语句)
2. (2分) 引入非终结符 M 及其 ε 产生式, 并在上述翻译模式的基础上添加如下语义计算规则:

- (6) $M \rightarrow \varepsilon \quad \{ M.s := 0 \}$

试用 M 替代嵌在上述翻译模式产生式中间的非复写语义动作, 添加适当的复写规则, 使得嵌在产生式中间的语义动作集中仅含复写规则, 并使得在自底向上的语义计算过程中, 文法符号的所有继承属性均可以通过归约前已出现在分析栈中的确定的综合属性进行访问。(如有必要, 可增加新的非终结符和相应的产生式)

3. (4分) 如果在 LR 分析过程中根据这一修改后新的翻译模式进行自下而上翻译, 试写出在按每个产生式归约时语义处理的一个代码片断 (设语义栈由向量 val 表示, 归约前栈顶位置为 top , 终结符不对应语义值, 而每个非终结符的综合属性都只对应一个语义值, 本题中可用 $val[i].s$ 表示; 不用考虑对 top 的维护)。
4. (5分) 文法 $G[S]$ 可用于识别后缀形式 (逆波兰式) 的命题表达式。输入串 $a b \neg \wedge a b c \wedge \vee a \neg c b \wedge \vee \vee \vee$ 对应于中缀式 $a \wedge \neg b \vee ((a \vee b \wedge c) \vee (\neg a \vee (c \wedge b)))$, 以下是该命题表达式对应的表达式树:



假设在一个简单的基于寄存器的机器 M 上进行表达式求值, 除了 `load/store` 指令用于寄存器值的装入和保存外, 其余操作均由下列格式的指令完成:

OP reg0, reg1, reg2

OP reg0, reg1

其中，reg0, reg1, reg2处可以是任意的寄存器，OP 为运算符。运行这些指令时，对 reg1和reg2的值做二元运算，或者对reg1的值做一元运算，结果存入reg0。对于 load/store指令，假设其格式为：

LD reg, mem /* 取内存或立即数 mem 的值到寄存器 reg */

ST reg, mem /* 存寄存器 reg 的值到内存量 mem */

我们假设M机器指令中，逻辑运算 \wedge 、 \vee 、 \neg 分别用助记符 AND、OR、NOT 表示。

试说明，为上述表达式树生成机器 M 指令序列时，需要寄存器书目的最小值 $n=?$ 假设这些寄存器分别用助记符 R_0, R_1, \dots ，和 R_{n-1} 表示，试采用课程中所介绍的方法生成该命题表达式的目标代码（仅含指令AND、OR、NOT、LD和ST，以及仅用寄存器 R_0, R_1, \dots ，和 R_{n-1} ）。（给出算法执行结果即可，不必进行目标代码优化）

七 (18分)

1. (6分) 以下是一个S-翻译模式片断，描述了某小语言部分特性的类型检查工作：

```
P → D ; S      { P.type := if D.type = ok and S.type = ok then ok else type_error }
S → S'         { S.type := S'.type }
S → if E then S1 { S.type := if E.type=bool then S1.type else type_error }
S → while E begin S1 end { S.type := if E.type=bool then S1.type else type_error }
S → for ( S'1; E ; S'2 ) begin S1 end
      { S.type := if E.type=bool and S'1.type = ok and S'2.type = ok
      then S1.type else type_error }
S → S1 ; S2    { S.type := if S1.type = ok and S2.type = ok then ok else type_error }
S' → id := E'   { S'.type := if lookup_type (id.entry) = E'.type then ok else type_error }
D → ...        /*省略与声明语句相关的全部规则*/
E → ...        /*省略与布尔表达式相关的全部规则*/
E' → ...       /*省略与算术表达式相关的全部规则*/
```

其中，type 属性以及类型表达式 *ok*, *type_error*, *bool*, 以及所涉及到的语义函数（如 *lookup_type*）等的含义与讲稿中一致；加黑的单词为保留字；声明语句、布尔表达式以及算术表达式相关的部分已全部略去。

（此外，假设在语法制导语义处理的分析过程中遇到的冲突问题均可以按照某种原则处理，我们不考虑基础文法是否 LR 文法。）

下面叙述本小题的要求：

若在基础文法中增加如下产生式，用于定义针对for-循环的“跳过本次循环体”语句：

$S \rightarrow \text{leap}$

语句 *leap* 只能出现在for-循环语句内部，但同时不能直接出现在某个while-循环语句

的内部。以下代码片断中，列举了几个合法与不合法使用leap语句的例子：

```

...          /*不含循环语句*/
leap;         /*此处的leap是不合法的*/
...          /*不含循环语句*/
for (i:=1; i<100; i:=i+1)
begin
  ...
  leap;       /*此处的leap是合法的*/
  ...
  if i=50 then leap;    /*此处的leap是合法的*/
  ...
  while cond
  begin
    ...
    leap;       /*此处的leap是不合法的*/
    ...
    for (...)
    begin
      leap;     /*此处的leap是合法的*/
    end
    ...
  end
end

end
...

```

试在上述翻译模式片段基础上增加相应的语义处理内容（要求是 *L*-翻译模式），以实现针对 `leap` 语句的类型检查（合法性检查）任务。

（提示：可以引入 *S* 的一个继承属性）

注：未修改的规则可省略不写。每条规则中，可仅给出与变化内容相关的规则。

2. (6分) 以下是一个*L*-翻译模式片断，可以产生相应的 *TAC* 语句序列：

$$\begin{aligned}
 P &\rightarrow D ; \{ S.next := newlabel \} S \{ gen(S.next ':') \} \\
 S &\rightarrow \{ S'.next := S.next \} S' \{ S.code := S'.code \} \\
 S &\rightarrow \text{if } \{ E.true := newlabel; E.false := S.next \} E \text{ then} \\
 &\quad \{ S_1.next := S.next \} S_1 \{ S.code := E.code // gen(E.true ':') // S_1.code \} \\
 S &\rightarrow \text{while } \{ E.true := newlabel; E.false := S.next \} E \\
 &\quad \text{begin } \{ S_1.next := newlabel \} S_1 \text{ end} \\
 &\quad \{ S.code := gen(S_1.next ':') // E.code // gen(E.true ':') // \\
 &\quad \quad S_1.code // gen('goto' S_1.next) \} \\
 S &\rightarrow \text{for } (\{ S'_1.next := newlabel \} S'_1; \\
 &\quad \{ E.true := newlabel; E.false := S.next \} E ; \\
 &\quad \{ S'_2.next := S'_1.next \} S'_2) \\
 &\quad \text{begin } \{ S_1.next := newlabel \} S_1 \text{ end } \{ \\
 &\quad \quad S.code := S'_1.code // gen(S'_1.next ':') // E.code // \\
 &\quad \quad gen(E.true ':') // S_1.code // gen(S_1.next ':') // \\
 &\quad \quad S'_2.code // gen('goto' S'_1.next) \} \\
 S &\rightarrow \{ S_1.next := newlabel \} S_1 ; \\
 &\quad \{ S_2.next := S.next \} S_2 \\
 &\quad \{ S.code := S_1.code // gen(S_1.next ':') // S_2.code \} \\
 S' &\rightarrow \underline{id} := E' \quad \{ S'.code := E'.code // gen(\underline{id}.place ':=' E'.place) \} \\
 D &\rightarrow \dots \quad /*省略与声明语句相关的全部规则*/
 \end{aligned}$$

$E \rightarrow \dots$ /*省略与布尔表达式相关的全部规则*/
 $E' \rightarrow \dots$ /*省略与算术表达式相关的全部规则*/

其中, 属性 $S.code, S'.code, E.code, S.next, S'.next, E.true, E.false$, 语义函数 $newlabel, gen()$ 以及所涉及到的 TAC 语句与讲稿中一致。语义函数 $newtemp$ 的作用是在符号表中新建一个从未使用过的名字, 并返回该名字的存储位置; 语义函数 gen 的结果是生成一条 TAC 语句; “||” 表示 TAC 语句序列的拼接。所有符号的 $place$ 综合属性也均与讲稿中一致。

(此外, 和前面一样, 假设在语法制导语义处理的分析过程中遇到的冲突问题均可按照某种原则处理, 我们不考虑基础文法是否 LR 文法。)

下面叙述本小题的要求:

若在基础文法中增加如下产生式, 用于定义针对for-循环的“跳过本次循环体”语句:

$S \rightarrow \text{leap}$

如第1小题中所述, 语句 leap 只能出现在for-循环语句内部, 但同时不能直接出现在某个while-循环语句的内部, 其执行语义为: 设 $\text{for}(S'_1; E; S'_2) \text{begin} \dots \text{end}$ 是直接包围该 leap 语句的 for-循环, 在执行 leap 语句后, 控制将跳过该循环体内部剩余的语句, 跳转到下一次循环 (先执行 S'_2)。

试在上述 L-翻译模式片段基础上增加针对 leap 语句的语义处理内容 (不改变 L-翻译模式的特征, 不考虑第1小题中已完成的语义检查工作)。

注: 可引入新的属性或删除旧的属性, 必要时给出解释。未修改的规则可省略不写。

3. (6分) 以下是一个 S-翻译模式片断, 可以产生相应的 TAC 语句序列:

```

 $P \rightarrow D; S M$           {  $\text{backpatch}(S.\text{nextlist}, M.\text{gotostm})$  }
 $S \rightarrow S'$               {  $S.\text{nextlist} := ""$  }
 $S \rightarrow \text{if } E \text{ then } M_1 S_1 N \text{ else } M_2 S_2$  {  $\text{backpatch}(E.\text{truelist}, M_1.\text{gotostm});$ 
                                      $\text{backpatch}(E.\text{falselist}, M_2.\text{gotostm});$ 
                                      $S.\text{nextlist} := \text{merge}(S_1.\text{nextlist}, \text{merge}(N.\text{nextlist}, S_2.\text{nextlist}))$  }
 $S \rightarrow \text{while } M_1 E \text{ begin } M_2 S_1 N \text{ end}$  {  $\text{backpatch}(S_1.\text{nextlist}, M_1.\text{gotostm});$ 
                                      $\text{backpatch}(E.\text{truelist}, M_2.\text{gotostm});$ 
                                      $\text{backpatch}(N.\text{nextlist}, M_1.\text{gotostm});$ 
                                      $S.\text{nextlist} := E.\text{falselist};$  }
 $S \rightarrow \text{for}(S'_1; M_1 E; M_2 S'_2 N_1) \text{ begin } M_3 S_1 N_2 \text{ end}$ 
                                     {  $\text{backpatch}(E.\text{truelist}, M_3.\text{gotostm});$ 
                                      $\text{backpatch}(S_1.\text{nextlist}, M_2.\text{gotostm});$ 
                                      $\text{backpatch}(N_1.\text{nextlist}, M_1.\text{gotostm});$ 
                                      $\text{backpatch}(N_2.\text{nextlist}, M_2.\text{gotostm});$ 
                                      $S.\text{nextlist} := E.\text{falselist}$  }
 $S \rightarrow S_1; M S_2$       {  $\text{backpatch}(S_1.\text{nextlist}, M.\text{gotostm});$ 
                                      $S.\text{nextlist} := S_2.\text{nextlist}$  }
 $S' \rightarrow \text{id} := E'$     {  $\text{emit}(\text{id}.\text{place} := E'.\text{place})$  }
 $M \rightarrow \varepsilon$            {  $M.\text{gotostm} := \text{nextstm}$  }
 $N \rightarrow \varepsilon$            {  $N.\text{nextlist} := \text{makelist}(\text{nextstm}); \text{emit}('goto \_')$  }
 $D \rightarrow \dots$           /*省略与声明语句相关的全部规则*/
 $E \rightarrow \dots$           /*省略与布尔表达式相关的全部规则*/
 $E' \rightarrow \dots$          /*省略与算术表达式相关的全部规则*/

```

其中, 所用到的属性和语义函数与讲稿中一致: 综合属性 $E.\text{truelist}$ (真链), 表示一系列跳转语句的地址, 这些跳转语句的目标语句标号是体现布尔表达式 E 为“真”的标号; 综合属性 $E.\text{falselist}$ (假链), 表示一系列跳转语句的地址, 这些跳转语句的

目标语句标号是体现布尔表达式 E 为“假”的标号；综合属性 $S.nextlist$ ($next$ 链), 表示一系列跳转语句的地址, 这些跳转语句的目标语句标号是在执行序列中紧跟在 S 之后的下条 TAC 语句的标号；语义函数 $makelist(i)$, 用于创建只有一个结点 i 的表, 对应于一条跳转语句的地址；语义函数 $merge(p_1, p_2)$, 表示连接两个链表 p_1 和 p_2 , 返回结果链表；语义函数 $backpatch(p, i)$, 表示将链表 p 中每个元素所指向的跳转语句的标号置为 i ；语义函数 $nextstm$, 将返回下一条 TAC 语句的地址；语义函数 $emit(...)$, 将输出一条 TAC 语句, 并使 $nextstm$ 加1。同第2小题, 所有符号的 $place$ 综合属性也均与讲稿中一致。

(和前面一样, 假设在语法制导语义处理的分析过程中遇到的冲突问题均可以按照某种原则处理, 这里不必考虑基础文法是否 LR 文法。)

下面叙述本小题的要求:

若在基础文法中增加如下产生式, 用于定义针对for-循环的“跳过本次循环体”语句:

$$S \rightarrow \text{leap}$$

如第1小题中所述, 语句 leap 只能出现在for-循环语句内部, 但同时不能直接出现在某个while-循环语句的内部, 其执行语义为: 设 $\text{for}(S'_1; E; S'_2) \text{begin} \dots \text{end}$ 是直接包围该 leap 语句的 for-循环, 在执行 leap 语句后, 控制将跳过该循环体内部剩余的语句, 跳转到下一次循环 (先执行 S'_2)。

试在上述 S -翻译模式片段基础上增加针对 leap 语句的语义处理内容 (不改变 S -翻译模式的特征, 不考虑第1小题中已完成的语义检查工作)。

注: 可引入新的属性或删除旧的属性, 必要时给出解释。未修改的规则可省略不写。

