

第一章 习题解答

1. 解：源程序是指以某种程序设计语言所编写的程序。目标程序是指编译程序（或解释程序）将源程序处理加工而得的另一种语言（目标语言）的程序。翻译程序是将某种语言翻译成另一种语言的程序的统称。编译程序与解释程序均为翻译程序，但二者工作方法不同。解释程序的特点是并不先将高级语言程序全部翻译成机器代码，而是每读入一条高级语言程序语句，就用解释程序将其翻译成一段机器指令并执行之，然后再读入下一条语句继续进行解释、执行，如此反复。即边解释边执行，翻译所得的指令序列并不保存。编译程序的特点是先将高级语言程序翻译成机器语言程序，将其保存到指定的空间中，在用户需要时再执行之。即先翻译、后执行。
2. 解：一般说来，编译程序主要由词法分析程序、语法分析程序、语义分析程序、中间代码生成程序、代码优化程序、目标代码生成程序、信息表管理程序、错误检查处理程序组成。
3. 解：C 语言的关键字有：auto break case char const continue default do double else enum extern float for goto if int long register return short signed sizeof static struct switch typedef union unsigned void volatile while。上述关键字在 C 语言中均为保留字。
4. 解：C 语言中括号有三种：{}，[]，()。其中，{}用于语句括号；[]用于数组；()用于函数（定义与调用）及表达式运算（改变运算顺序）。C 语言中无 END 关键字。逗号在 C 语言中被视为分隔符和运算符，作为优先级最低的运算符，运算结果为逗号表达式最右侧子表达式的值（如：(a, b, c, d) 的值为 d）。
5. 略

第二章 习题解答

1. (1) 答: $26*26=676$

(2) 答: $26*10=260$

(3) 答: $\{a, b, c, \dots, z, a0, a1, \dots, a9, aa, \dots, az, \dots, zz, a00, a01, \dots, zzz\}$,
共 $26+26*36+26*36*36=34658$ 个

2. 构造产生下列语言的文法

(1) $\{anbn \mid n \geq 0\}$

解: 对应文法为 $G(S) = (\{S\}, \{a, b\}, \{S \rightarrow \epsilon \mid aSb\}, S)$

(2) $\{anbmcp \mid n, m, p \geq 0\}$

解: 对应文法为 $G(S) = (\{S, X, Y\}, \{a, b, c\}, \{S \rightarrow aS \mid X, X \rightarrow bX \mid Y, Y \rightarrow cY \mid \epsilon\}, S)$

(3) $\{an \# bn \mid n \geq 0\} \cup \{cn \# dn \mid n \geq 0\}$

解: 对应文法为 $G(S) = (\{S, X, Y\}, \{a, b, c, d, \#\}, \{S \rightarrow X, S \rightarrow Y, X \rightarrow aXb \mid \#, Y \rightarrow cYd \mid \#\}, S)$

(4) $\{w\#wr\# \mid w \in \{0, 1\}^*\text{, wr 是 w 的逆序排列}\}$

解: $G(S) = (\{S, W, R\}, \{0, 1, \#\}, \{S \rightarrow W\#, W \rightarrow 0W0 \mid 1W1 \mid \#\}, S)$

(5) 任何不是以 0 打头的所有奇整数所组成的集合

解: $G(S) = (\{S, A, B, I, J\}, \{-, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}, \{S \rightarrow J \mid IBJ, B \rightarrow 0B \mid IB \mid e, I \rightarrow J \mid 2 \mid 4 \mid 6 \mid 8, J \rightarrow 1 \mid 3 \mid 5 \mid 7 \mid 9\}, S)$

(6) 所有偶数个 0 和偶数个 1 所组成的符号串集合

解: 对应文法为 $S \rightarrow 0A \mid 1B \mid e, A \rightarrow 0S \mid 1C, B \rightarrow 0C \mid 1S, C \rightarrow 1A \mid 0B$

3. 描述语言特点

(1) $S \rightarrow 10S0 \mid aAA \rightarrow bAA \rightarrow a$

解: 本文法构成的语言集为: $L(G) = \{(10)^n ab^m a^0 \mid n, m \geq 0\}$ 。

(2) $S \rightarrow SS \quad S \rightarrow 1A0A \rightarrow 1A0A \rightarrow \epsilon$

解: $L(G) = \{1n10n11n20n2 \dots 1nm0nm \mid n_1, n_2, \dots, n_m \geq 0\}$; 且 n_1, n_2, \dots, n_m 不全为零} 该语言特点是: 产生的句子中, 0、1 个数相同, 并且若干相接的 1 后必然紧接数量相同连续的 0。

(3) $S \rightarrow 1AS \rightarrow B0A \rightarrow 1AA \rightarrow CB \rightarrow B0B \rightarrow CC \rightarrow 1C0C \rightarrow \epsilon$

解: 本文法构成的语言集为:

$L(G) = \{1p1n0n \mid p \geq 1, n \geq 0\} \cup \{1n0n0q \mid q \geq 1, n \geq 0\}$, 特点是具有 $1p1n0n$ 或 $1n0n0q$ 形式, 进一步, 可知其具有形式 $1n0mn, m \geq 0$, 且 $n+m > 0$ 。

(4) $S \rightarrow bAdcA \rightarrow AGSG \rightarrow \epsilon \quad A \rightarrow a$

解: 可知, $S \Rightarrow \dots \Rightarrow baSndc \quad n \geq 0$

该语言特点是: 产生的句子中, 是以 ba 开头 dc 结尾的串, 且 ba、dc 个数相同。

(5) $S \rightarrow aSSS \rightarrow a$

解: $L(G) = \{a(2n-1) \mid n \geq 1\}$ 可知: 奇数个 a

4. 解: 此文法产生的语言是: 以终结符 $a_1, a_2 \dots a_n$ 为运算对象, 以 \wedge, \vee, \sim 为运算符, 以 $[,]$ 为分隔符的布尔表达式串

5. 5.1 解: 由于此文法包含以下规则: $AA \rightarrow e$, 所以此文法是 0 型文法。

5.2 证明: 略

6. 解:

(1) 最左推导:

$\langle \text{程序} \rangle T \langle \text{分程序} \rangle T \langle \text{标号} \rangle: \langle \text{分程序} \rangle TL: \langle \text{分程序} \rangle$

$TL: \langle \text{标号} \rangle: \langle \text{分程序} \rangle$

$T L: L: \langle \text{分程序} \rangle$

$T L: L: \langle \text{无标号分程序} \rangle$

$T L: L: \langle \text{分程序首部} \rangle; \langle \text{复合尾部} \rangle$

$T L: L: \langle \text{分程序首部} \rangle; \langle \text{说明} \rangle; \langle \text{复合尾部} \rangle$

T L: L: begin<说明>; <说明>; <复合尾部>

T L: L: begin d; <说明>; <复合尾部>

T L: L: begin d; d; <复合尾部>

T L: L: begin d; d; <语句>; <复合尾部>

T L: L: begin d; d; s; <复合尾部>

T L: L: begin d; d; s; <语句> end

T L: L: begin d; d; s; s end

最右推导:

<程序>T<分程序>T<标号>; <分程序>

T<标号>; <标号>; <分程序>

T<标号>; <标号>; <无标号分程序>

T<标号>; <标号>; <分程序首部>; <复合尾部>

T<标号>; <标号>; <分程序首部>; <语句>; <复合尾部>

T<标号>; <标号>; <分程序首部>; <语句>; <语句>; end

T<标号>; <标号>; <分程序首部>; <语句>; s; end

T<标号>; <标号>; <分程序首部>; s; s; end

T<标号>; <标号>; <分程序首部>; 说明; s; s; end

T<标号>; <标号>; <分程序首部>; d; s; s; end

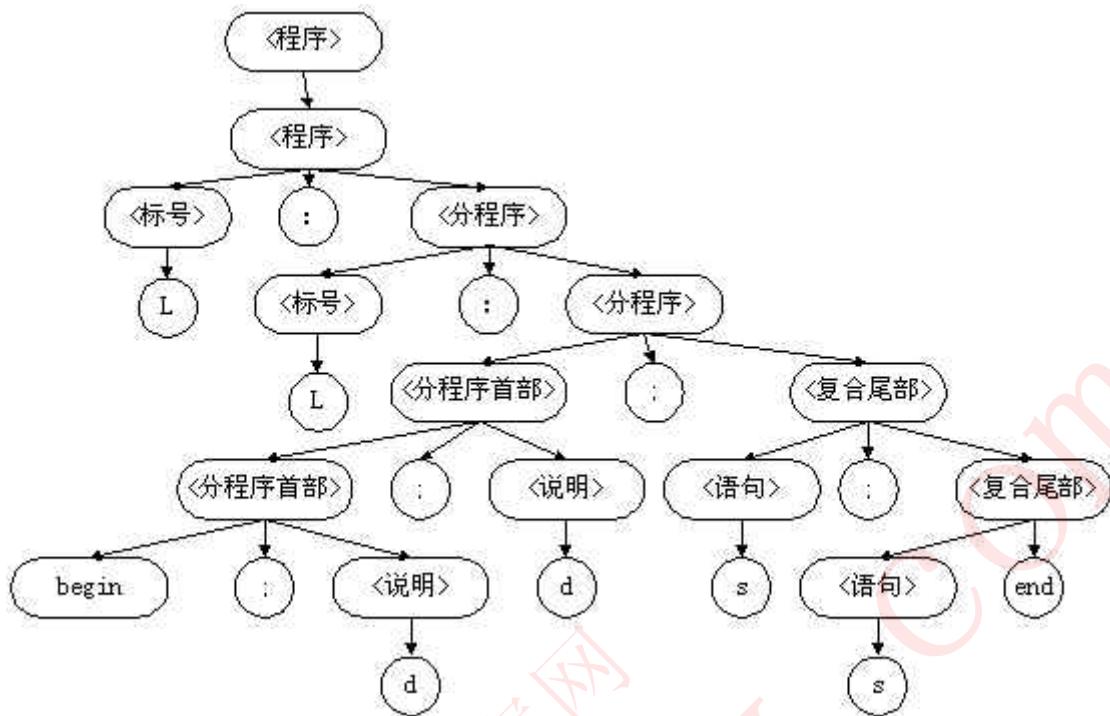
T<标号>; <标号>; begin 说明; d; s; s; end

T<标号>; <标号>; begin d; d; s; s; end

T<标号>; L: begin d; d; s; s; end

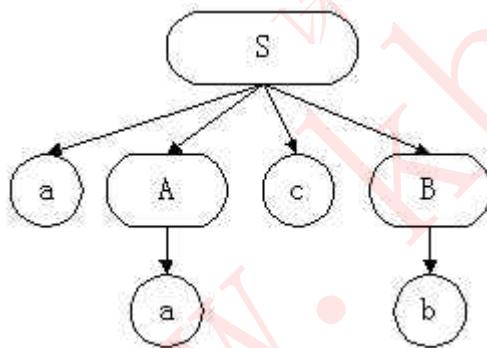
TL: L: begin d; d; s; s; end

(2) 句子 L: L: begin d; d; s; s end 的相应语法树是:



7. 解：

aacb 是文法 $G[S]$ 中的句子，相应语法树是：



最右推导： $S \Rightarrow aAcB \Rightarrow aAcB \Rightarrow aacb$

最左推导： $S \Rightarrow aAcB \Rightarrow aacB \Rightarrow aacb$

(2) aabacbadcd 不是文法 $G[S]$ 中的句子

因为文法中的句子不可能以非终结符 d 结尾

(3) aacbccb 不是文法 $G[S]$ 中的句子

可知， aacbccb 仅是文法 $G[S]$ 的一个句型的一部分，而不是一个句子。

(4) aacabcbcccaacdca 不是文法 $G[S]$ 中的句子

因为终结符 d 后必然要跟终结符 a, 所以不可能出现…dc…这样的句子。

(5) aacabcbcccaacbca 不是文法 $G[S]$ 中的句子

由 (1) 可知: aacb 可归约为 S, 由文法的产生式规则可知, 终结符 c 后不可能跟非终结符 S, 所以不可能出现…caacb…这样的句子。

8. 证明: 用归纳法于 $n, n=1$ 时, 结论显然成立。设 $n=k$ 时, 对于 $\alpha_1 \alpha_2 \dots \alpha_k T^* b$, 存在 $\beta_i: i=1, 2, \dots, k$, $\alpha_i T^* \beta_i$ 成立, 现在设

$\alpha_1 \alpha_2 \dots \alpha_k \alpha_{k+1} T^* b$, 因文法是前后文无关的, 所以 $\alpha_1 \alpha_2 \dots \alpha_k$ 可推导出 b 的一个前缀 b' , α_{k+1} 可推导出 b 的一个后缀 $=b''$ (不妨称为 b_{k+1})。由归纳假设, 对于 b' , 存在 $\beta_i: i=1, 2, \dots, k$, $b' = \beta_1 \beta_2 \dots \beta_k$, 使得

$\alpha_i T^* \beta_i$ 成立, 另外, 我们有 $\alpha_{k+1} T^* b'' (=b_{k+1})$ 。即 $n=k+1$ 时亦成立。证毕。

9. 证明: (1) 用反证法。假设 α 首符号为终结符时, β 的首符号为非终结符。即设: $\alpha = a \omega$; $\beta = A \omega'$ 且 $\alpha \Rightarrow^* \beta$ 。

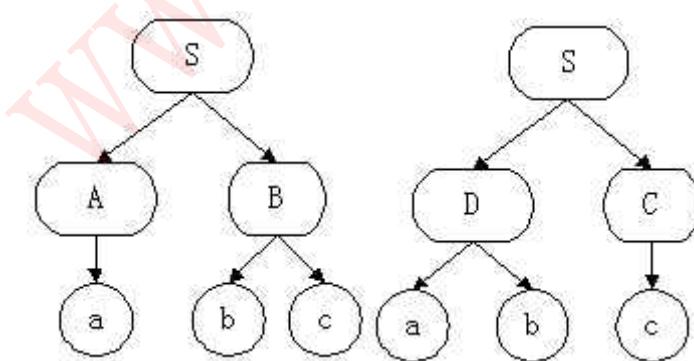
由题意可知: $\alpha = a \omega T \dots T A \omega' = \beta$, 由于文法是 CFG, 终结符 a 不可能被替换为空串或非终结符, 因此假设错误。得证;

(2) 同 (1), 假设: β 的首符号为非终结符时, α 首符号为终结符。即设: $\alpha = a \omega$; $\beta = A \omega'$ 且 $\alpha = a \omega T \dots T A \omega' = \beta$, 与 (1) 同理, 得证。

10. 证明: 因为存在句子: abc, 它对应有两个语法树 (或最右推导):

STABTAbcTabc

STDCTDcTabc

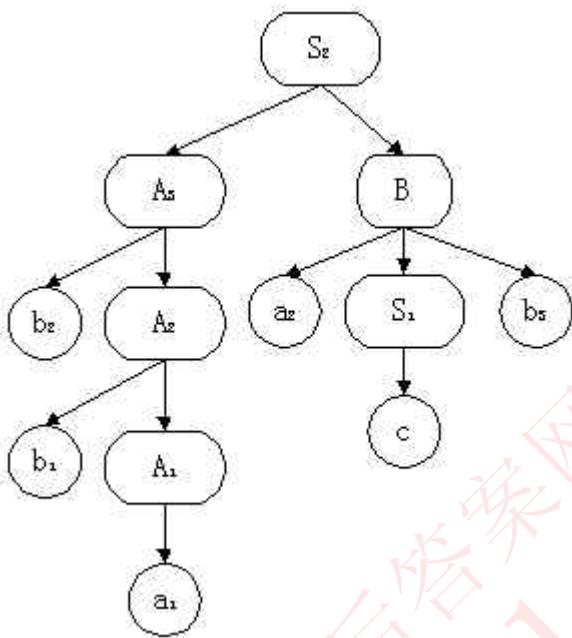


所以, 本文法具有二义性。

11. 解：

(1) STABTA~~a~~S~~b~~TAacbT~~b~~AacbT~~bb~~AacbT~~bb~~aacb

上面推导中，下划线部分为当前句型的句柄。对应的语法树为：



全部的短语：

第一个 a (a1) 是句子 bbaacb 相对于非终结符 A (A1) (产生式 A?a) 的短语 (直接短语)；

b1a1 是句子 bbaacb 相对于非终结符 A2 的短语；

b2b1a1 是句子 bbaacb 相对于非终结符 A3 的短语；

c 是句子 bbaacb 相对于非终结符 S1 (产生式 S?c) 的短语 (直接短语)；

a2cb3 是句子 bbaacb 相对于非终结符 B 的短语；

b2b1a1a2cb3 是句子 bbaacb 相对于非终结符 S2 的短语；

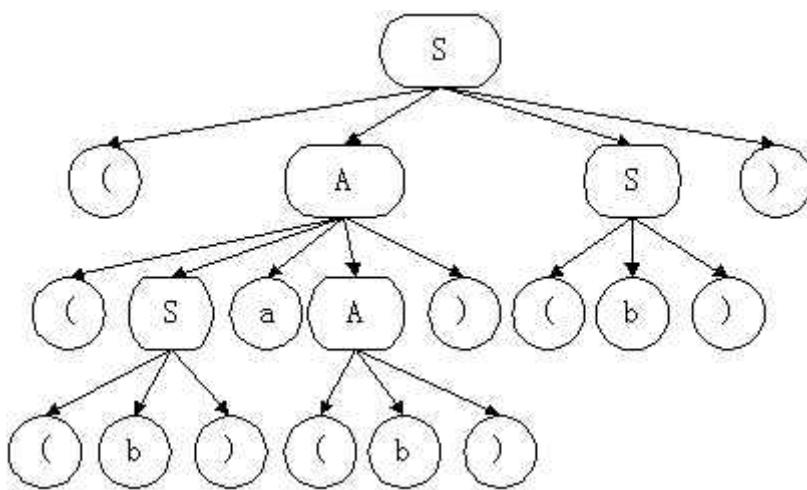
注：符号的下标是为了描述方便加上去的。

(2) 句子 ((b) a (a)) (b)) 的最右推导：

ST (AS) T (A (b)) T ((SaA) (b)) T ((Sa (a)) (b))

T (((b) a (a)) (b))

相应的语法树是：



(3) 解: $i \rightarrow i * i \rightarrow i + i$ 对应的语法树略。

最右推导: $E \rightarrow TT \rightarrow FP \uparrow T \rightarrow FE \uparrow T \rightarrow FET+ \uparrow T \rightarrow FEF+ \uparrow T \rightarrow FEP+ \uparrow T \rightarrow FEi+ \uparrow T$

TFTi+ ↑ T FTF*i+ ↑ TFTP*i+ ↑ T FTi*i+ ↑ TFFi*i+ ↑ T FPi*i+ ↑

TF_i i*i+↑ T P_i i*i+↑ T_i i*i+↑

12. 证明: 公式 (1) 与 (2) 的等价性. 当前文涉及的每一个符号中仅有一个包打和一个包打产生式且当前符号

必要性：有唯一的语法树 T 对每一步推导都有唯一的最右推导 T 对当前符号串有

唯一的最近归约 1 之前

18. 亂世有才者不文法

15. 消除下列文法中的无用产生式和单产生式

(1) 消除后的产生式如下:

$S \rightarrow aB \mid BC$

$B \rightarrow DB \mid b$

$C \rightarrow b$

$D \rightarrow b \mid DB$

(2) 消除后的产生式如下:

$S \rightarrow SA \mid SB \mid () \mid (S) \mid [] \mid [S]$

$A \rightarrow () \mid (S) \mid [] \mid [S]$

$B \rightarrow [] \mid [S]$

(3) 消除后的产生式如下:

$E \rightarrow E+T \mid T*F \mid (E) \mid P \uparrow F \mid i$

$T \rightarrow T*F \mid (E) \mid P \uparrow F \mid i$

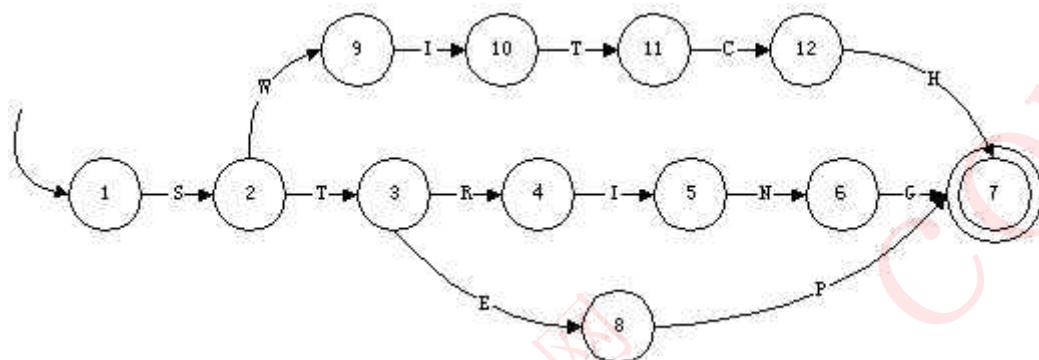
$F \rightarrow P \uparrow F \mid (E) \mid i$

$P \rightarrow (E) \mid i$

第三章 习题解答

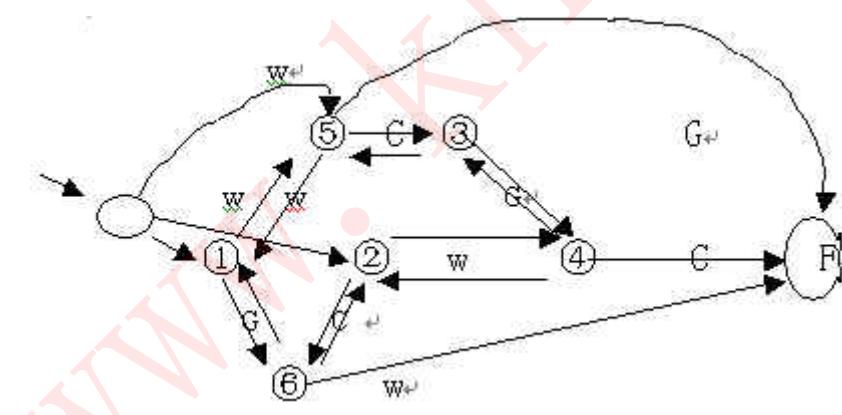
1. 从略

2.



3 假设 W: 表示载狐狸过河, G: 表示载山羊过河, C: 表示载白菜过河

用到的状态 1: 狐狸和山羊在左岸 2: 狐狸和白菜载左岸 3: 羊和白菜在左岸 4: 狐狸和山羊在右岸 5: 狐狸和白菜在右岸 6: 山羊和白菜在右岸 F: 全在右岸



4 证明: 只须证明文法 $G: A \rightarrow aB$ 或 $A \rightarrow a$ ($A, B \in VN, a \in VT^+$)

等价于 $G1: A \rightarrow aB$ 或 $A \rightarrow a$ ($a \in VT^+$)

- $G1$ 的产生式中 $A \rightarrow aB$, 则 B 也有 $B \rightarrow bC, C \rightarrow cD \dots$

所以有 $A \rightarrow abc\dots B'$, $a, b, c \dots \in VT, B' \in VN$

所以与 G 等价。

2) G 的产生式 $A \rightarrow \alpha B$, $\alpha \in VT^+$, 因为 α 是字符串, 所以肯定存在着一个终结符 a , 使 $A \rightarrow aB$

可见两者等价, 所以由此文法产生的语言是正规语言。

5

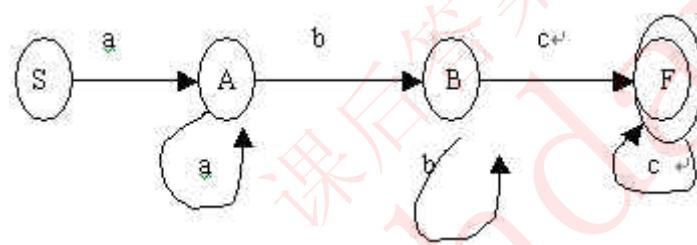
6 根据文法知其产生的语言是

$$L = \{ambnci \mid m, n, i \geq 1\}$$

可以构造如下的文法 $VN = \{S, A, B, C\}$, $VT = \{a, b, c\}$

$$P = \{ S \rightarrow aA, A \rightarrow aA, A \rightarrow bB, B \rightarrow bB, B \rightarrow cC, C \rightarrow cC, C \rightarrow c \}$$

其状态转换图如下:



7 (1) 其对应的右线性文法是:

$$A \rightarrow 0D, B \rightarrow 0A, B \rightarrow 1C, C \rightarrow 1 \mid 1F, C \rightarrow 1 \mid 0A, F \rightarrow 0 \mid 0E \mid 1A, D \rightarrow 0B \mid 1C, E \rightarrow 1C \mid 0B$$

(2) 最短输入串 011

(3) 任意接受的四个串

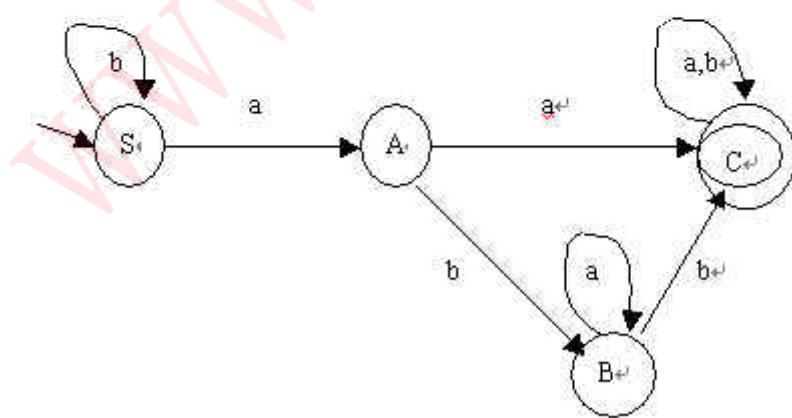
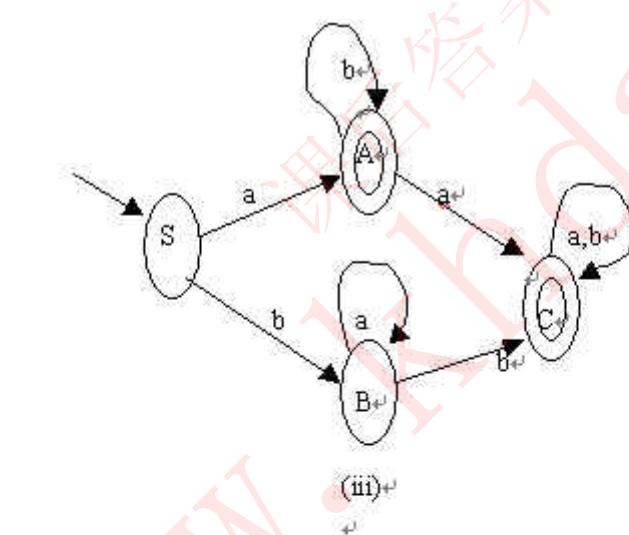
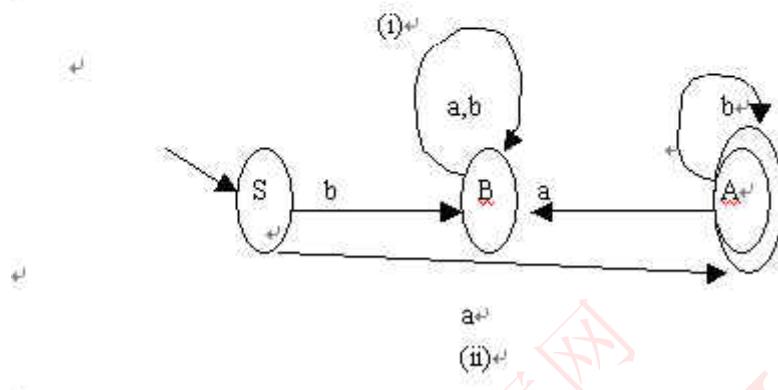
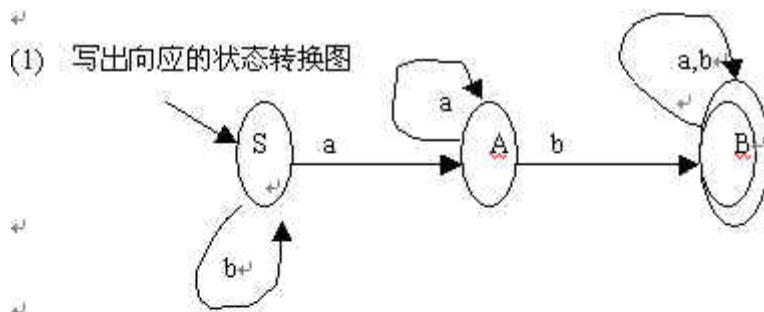
011, 0110, 0011, 000011

(4) 任意以 1 打头的串.

8 从略。

9

(1) 写出向应的状态转换图



(iv)

(2) 相应的 3 型文法

- (i) $S \rightarrow aAS \rightarrow bS \quad A \rightarrow aA \quad A \rightarrow bB \quad B \rightarrow a \mid aB \quad B \rightarrow b \mid bB$
- (ii) $S \rightarrow aA \mid a \quad S \rightarrow bB \quad B \rightarrow aB \mid bB \quad A \rightarrow aB \quad A \rightarrow b \mid bA$
- (iii) $S \rightarrow aA \quad S \rightarrow bB \quad A \rightarrow bA \quad A \rightarrow aC \quad B \rightarrow aB \quad B \rightarrow bC \quad C \rightarrow a \mid aC \quad C \rightarrow b \mid bC$
- (iv) $S \rightarrow bS \quad S \rightarrow aA \quad A \rightarrow aC \quad A \rightarrow bB \quad B \rightarrow aB \quad B \rightarrow bC \quad C \rightarrow a \mid aC \quad C \rightarrow b \mid bC$

(3) 用自然语言描述输入串的特征

(i) 以任意个(包括 0) b 开头, 中间有任意个(大于 1) a, 跟一个 b, 还可以有一个由 a, b 组成的任意字符串

(ii) 以 a 打头, 后跟任意个(包括 0) b

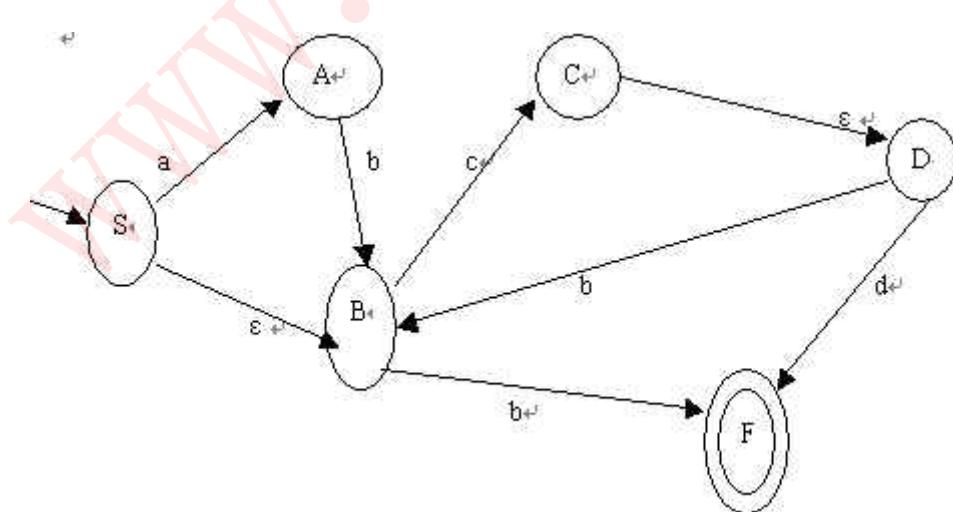
(iii) 以 a 打头, 中间有任意个(包括 0) b, 再跟 a, 最后由一个 a, b 所组成的任意串结尾或者

以 b 打头, 中间有任意个(包括 0) a, 再跟 b, 最后由一个 a, b 所组成的任意串结尾

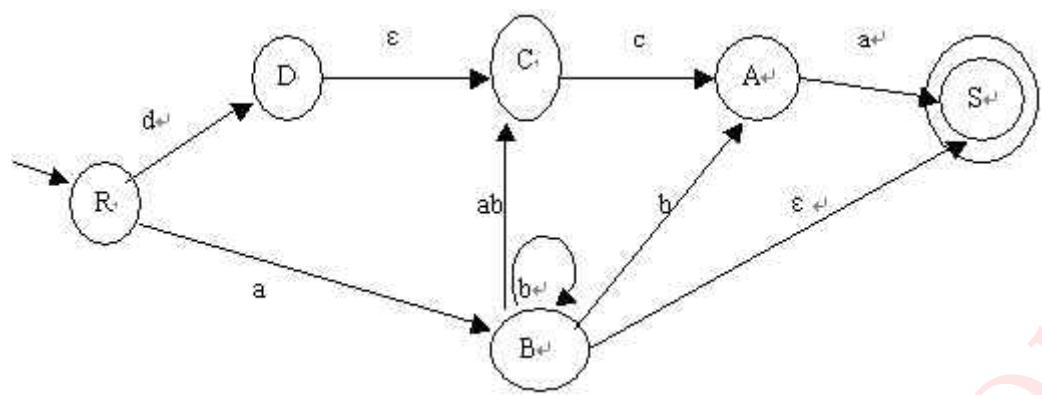
(iv) 以任意个(包括 0) b 开头, 中间跟 aa 最后由一个 a, b 所组成的任意串结尾或者

以任意个(包括 0) b 开头, 中间跟 ab 后再接任意(包括 0) a 再接 b, 最后由一个 a, b 所组成的任意串结尾

10 (1) G1 的状态转换图:



G2 的状态转换图:



(2) G1 等价的左线性文法:

$S \rightarrow Bb, S \rightarrow Dd, D \rightarrow C, B \rightarrow Db, C \rightarrow Bc, B \rightarrow Ab, B \rightarrow \epsilon, A \rightarrow a$

G2 等价的右线性文法:

$S \rightarrow dD, S \rightarrow aB, D \rightarrow C, B \rightarrow abC, B \rightarrow bB, B \rightarrow bA, B \rightarrow \epsilon, C \rightarrow cA, A \rightarrow a$

(3) 对 G1 文法, abb 的推导序列是:

$S \Rightarrow aA \Rightarrow abB \Rightarrow abb$

对 G1' 文法, abb 的推导序列是:

$S \Rightarrow Bb \Rightarrow Abb \Rightarrow abb$

对 G2 文法, aabca 的推导序列是:

$S \Rightarrow Aa \Rightarrow Cca \Rightarrow Babca \Rightarrow aabca$

对 G2' 文法, aabca 的推导序列是:

$S \Rightarrow aB \Rightarrow aabC \Rightarrow aabcA \Rightarrow aabca$

(4) 对串 acbd 来说, G1, G1' 文法都不能产生。

11 将右线性文法化为左线性文法的算法:

- (1) 对于 G 中每一个形如 $A \rightarrow aB$ 的产生式且 A 是开始符, 将其变为 $B \rightarrow a$, 否则若 A 不是开始符, $B \rightarrow Aa$;

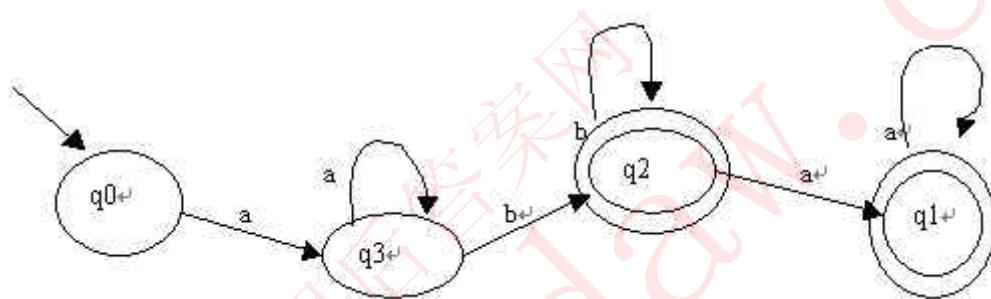
- (2) 对于 G 中每一个形如 $A \rightarrow a$ 的产生式, 将其变为 $S \rightarrow Aa$

12 (1)

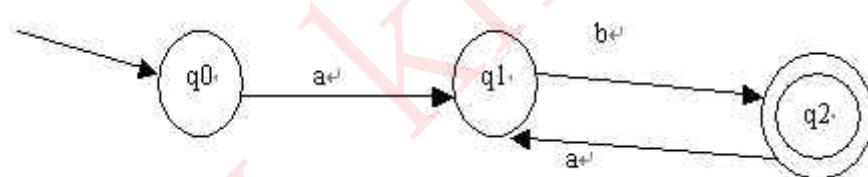
	a	b^ω
S	$\{S, A\}$	Φ^ω
A	Φ	$\{A, B\}^\omega$
B	$\{B\}$	Φ^ω

状态矩阵是:

记 $[S]=q_0$ $[B]=q_1$ $[A B]=q_2$ $[S A]=q_3$, 最小化和确定化后如图

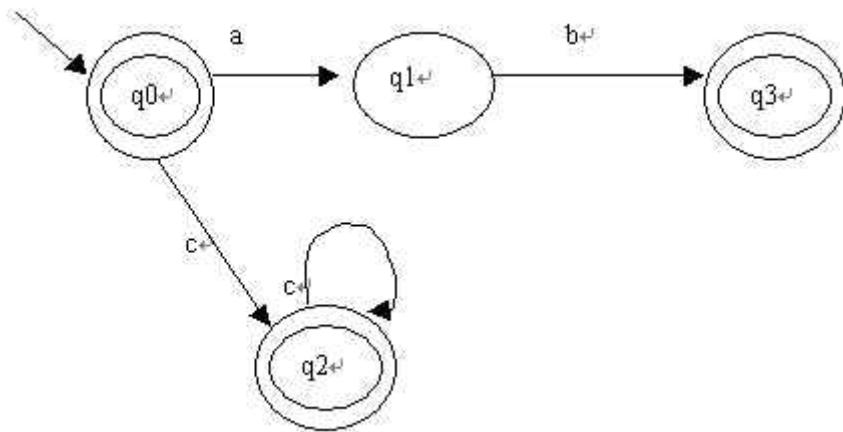


(2) 记 $[S]=q_0$, $[A]=q_1$, $[B S]=q_2$ 最小化和确定化后的状态转换图如下

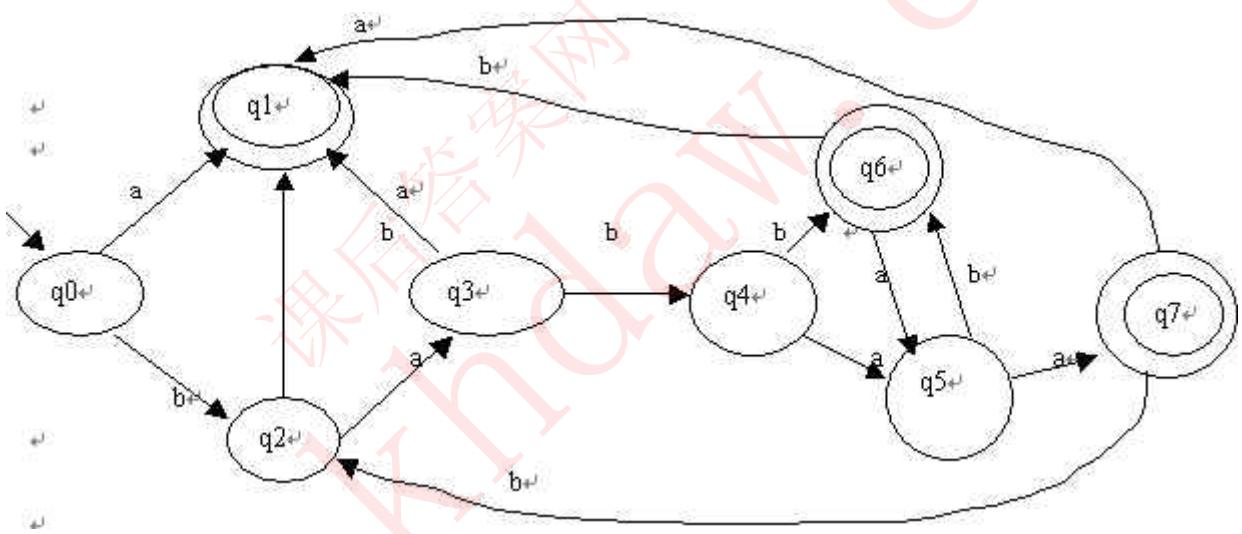


13 (1) 将具有 ϵ 动作的 NFA 确定化后, 其状态转换图如图:

记 $\{S_0, S_1, S_3\}=q_0$ $\{S_1\}=q_1$ $\{S_2, S_3\}=q_2$ $\{S_3\}=q_3$



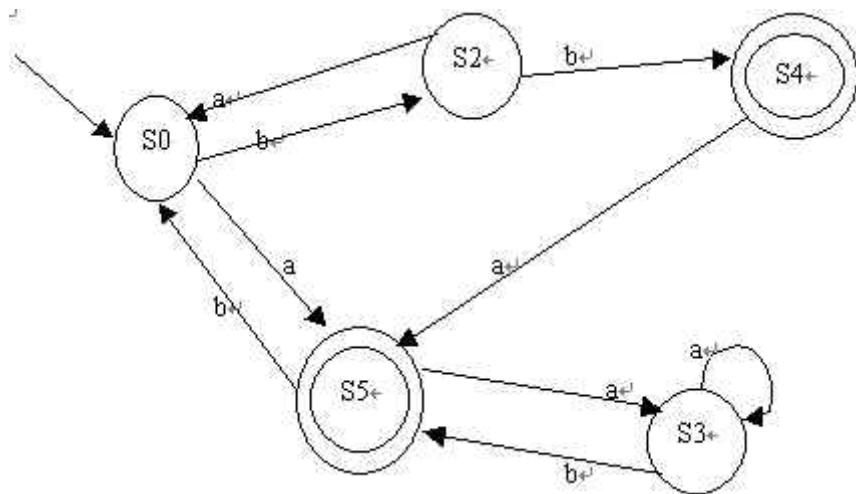
- (2) 记 $\{S\}=q_0$ $\{Z\}=q_1$ $\{U\ R\}=q_2$ $\{S\ X\}=q_3$ $\{Y\ U\ R\}=q_4$ $\{X\ S\ U\}=q_5$ $\{Y\ U\ R\ Z\}=q_6$ $\{Z\ S\}=q_7$



14 (1) 从略

(2) 化简后 S_0 和 S_1 作为一个状态, S_5 和 S_6 作为一个状态。

状态转换图如图



15 从略。

16 从略。

- (1) r^* 表示的正规式集是 $\{ \epsilon, r, rr, rrr, \dots \}$

$(\epsilon \mid r)^*$ 表示的正规式集是 $\{ \epsilon, \epsilon \mid \epsilon, \dots \} \cup \{ r, rr, rrr, \dots \} = \{ \epsilon, r, rr, rrr, \dots \}$

$\epsilon \mid rr^*$ 表示的正规式集是 $\{ \epsilon, r, rr, rrr, \dots \}$

$(r^*)^* = r^* = \{ \epsilon, r, rr, rrr, \dots \}$

所以四者是等价的。

(2) $(rs)^* r$ 表示的正规式集是 $\{ \epsilon, rs, rsrs, rsrsrs, \dots \} r$
 $= \{ r, rsr, rsrsr, rsrsrsr, \dots \}$

$r(sr)^*$ 表示的正规式集是 $r \{ \epsilon, sr, srsr, srsrsr, \dots \}$
 $= \{ r, rsr, rsrsr, rsrsrsr, \dots \}$

所以两者等价。

18 写成方程组

$$S = aT + aS \quad (1)$$

$$B = cB + c \quad (2)$$

$$T = bT + bB \quad (3)$$

所以 $B = c * cT = b * bc * c$

$$S = a * ab * bc * c$$

• G1:

$$S = aA + B \quad (1)$$

$$B = cC + b \quad (2)$$

$$A = abS + bB \quad (3)$$

$$C = D \quad (4)$$

$$D = bB + d \quad (5)$$

把 (4) (5) 代入 (2) , 得 $B = c(bB + d) + b = cbB + cd + b$ 得 $B = (cb) * (cd | b)$, 代入 (3) 得

$$A = abS + b(cb) * (cd | b) \text{ 把它打入 (1) 得}$$

$$S = a(abS + b(cb) * (cd | b)) + (cb) * (cd | b)$$

$$= aabS + ab(cb) * (cd | b) + (cb) * (cd | b)$$

$$= (aab) * (ab(cb) * (cd | b) | (cb) * (cd | b))$$

G2:

$$S = Aa + B \quad (1)$$

$$A = Cc + Bb \quad (2)$$

$$B = Bb + a \quad (3)$$

$$C = D + Bab \quad (4)$$

$$D = d \quad (5)$$

$$\text{可得 } D = dB = ab * C = ab * ab | b \\ A = (ab * ab | b) c + ab * b$$

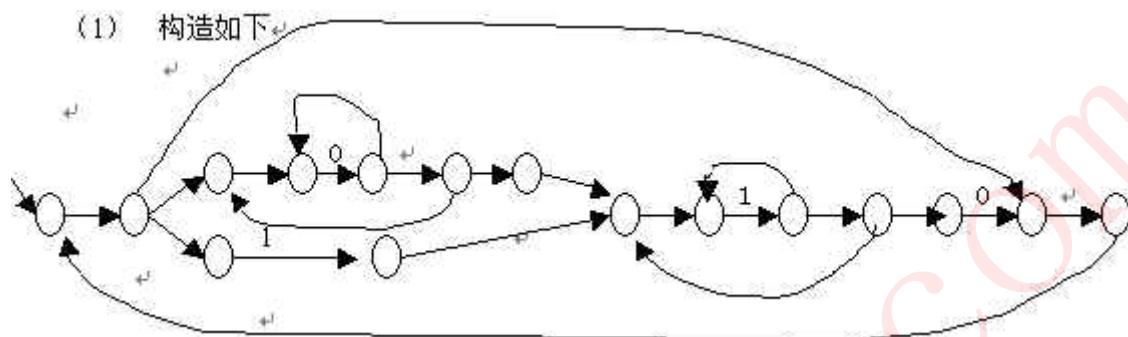
$$S = (ab * ab | b) ca + ab * ba + ab *$$

$$= (ab * ab | b) ca | ab * ba | ab *$$

- 识别此语言的正规式是 $S = 'LABEL' \ d(d|d)^*$;
- 从略。

21 从略。

22 构造 NFA



其余从略。

23 下面举一个能够识别 1, 2, 3, 10, 20, 100 的例子，读者可以推而广之。

```
%{  
#include <stdio.h>  
#include <string.h>  
#include <ctype.h>  
#define ON1  
#define TW 2  
#define THRE 3  
#define TE 10  
#define TWENT 20  
#define HUNDRE 100  
#define WHITE9999  
%}  
upper[A-Z]
```

```
%%

ONEreturn ON;

TWOreturn TW;

THREEreturn THRE;

TENreturn TE;

TWENTYreturn TWENT;

HUNDREDreturn HUNDRE;

" "+|\treturn WHITE;

\nreturn0;

%%

main(int argc, char *argv[])
{
    int c, i=0;
    char tmp[30];
    if (argc==2)
    {
        if ((yyin=fopen(argv[1], "r"))==NULL)
        {
            printf("can't open %s\n", argv[1]);exit(0);
        }
    }

    while ((c=yylex())!=0)
    {
```

```
switch(c)
{
    case 0N:
        c=yylex();
        if (c==0) goto {i+=1;label;}
        c=yylex();
        if (c==HUNDRE)
            i+=100;
        else i+=1;
        break;
    case TW:c=yylex();
        c=yylex();
        if (c==HUNDRE)
            i+=200;
        else i+=2;
        break;
    case TWENT: i+=20;
        break;
    case TE:i+=10;
        break;
    default:break;
}
}/*while*/
```

```
label: printf("%d\n", i);  
  
return;  
}
```

24 (1) D_n 表示的正规集是长度为 $2n$ 任意 a 和 b 组成的字符串。

- 此正规式的长度是 $2n$
- 用来识别 D_n 的 DFA 至多需要 $2n+1$ 个状态。

25 从略。

26(1)由 {} 括住的, 中间由任意个非 {} 组成的字符串, 如 {}, {{}}, {a}, {defg} 等等。

(2) 匹配一行仅由一个大写字母和一个数字组成的串, 如 A1, F8, Z2 等。

(3) 识别 \r\n 和除数字字符外的任何字符。

- 由 ' 和 ' 括住的, 中间由两个 ' 或者非 ' 和 \n 组成的任意次的字符串。如 ' , ' , ' , 'a' , ' bb' , ' def' , ' , ' , ' , ' 等等

270 [Xx] [0-9]* [a-fA-F]* | [0-9]+ | (\ ' ([a-zA-Z] | \\ [Xx] [0-7] [0-7a-fA-F] | \\ 0[01] [0-7] [0-7] | \\ [a-z]) \ ')

28 ^ [a-zA-Z_] + [0-9]* [a-zA-Z_] *

29 参考程序如下:

```
%{  
  
#include <stdio.h>  
  
#include <string.h>  
  
#include <ctype.h>  
  
#define UPPER2  
  
#define WHITE3  
  
%}  
  
upper[A-Z]
```

```
%%

{upper}+returnUPPER;

\t| " "+returnWHITE;

%%

main(int argc, char *argv[])
{
    int c, i;

    if (argc==2)
    {
        if ((yyin=fopen(argv[1], "r"))==NULL)
        {
            printf("can't open %s\n", argv[1]); exit(0);
        }
    }

    while ((c=yylex())!=EOF)
    {
        if (c==2)
        {
            for (i=0;yytext[i];i++)
                printf("%c", tolower(yytext[i]));
            yytext[0]=' \'000';
        }
        if (c==3)
```

```
printf(" ");  
else printf("%s", yytext);  
}  
  
return;  
}  
  
yywrap()  
{  
  
return ;  
}  
}
```

30 从略。

第四章 习题解答

1. 解：

(1) $S \rightarrow (S)Z21 \mid ()Z21 \mid [S]Z31 \mid []Z31$

$A \rightarrow (S)Z22 \mid ()Z22 \mid [S]Z32 \mid []Z32$

$B \rightarrow (S)Z23 \mid ()Z23 \mid [S]Z33 \mid []Z33$

$Z11 \rightarrow \epsilon \mid AZ11 \mid BZ21$

$Z12 \rightarrow AZ12 \mid BZ22 \mid Z13 \rightarrow AZ13 \mid BZ23$

$Z21 \rightarrow Z11Z22 \rightarrow \epsilon \mid Z12$

$Z23 \rightarrow Z13Z31 \rightarrow Z21$

$Z32 \rightarrow Z22Z33 \rightarrow \epsilon \mid Z23$

(2) $S \rightarrow bZ11 \mid aZ21 \mid A \rightarrow bZ12 \mid aZ22$

$Z11 \rightarrow \epsilon \mid AZ21 \mid Z12 \rightarrow AZ22 \mid Z21 \rightarrow SZ21 \mid Z22 \rightarrow \epsilon \mid SZ22$

(3) $S \rightarrow (T)Z11 \mid aZ11 \mid Z11 \mid T \rightarrow (T)Z12 \mid aZ12 \mid Z12$

$Z11 \rightarrow \epsilon \mid Z21 \mid Z12 \rightarrow Z22 \mid Z21 \rightarrow SZ21 \mid Z22 \rightarrow \epsilon \mid SZ22$

• 2. 解：

$S \Rightarrow AbB1, 1. 1$ (表示第 1 步, 用产生式 1. 1 推导, 以下同)

$\Rightarrow CABbB2, 2. 1$

$\Rightarrow edAbbB3, 4. 1$

$\Rightarrow edCAbbB4, 2. 1$

$\Rightarrow ededAbB5, 4. 1$

$\Rightarrow edaAbB5, 4. 2$ (不符合, 改写第 5 步, 用 4. 2)

$\Rightarrow edBfbbB4, 2. 2$

$\Rightarrow edCSdfbbB5, 3. 1$

⇒ededSdfbbB6, 4. 1

⇒edaSdfbbB6, 4. 2

⇒eddfbbB5, 3. 2

⇒eddfbbCSd6, 3. 1

⇒eddfbbbedSd7, 4. 1

⇒eddfbbaSd7, 4. 2

⇒eddfbbd6, 3. 2

- 3. 解：以下 Save 表示 save token_pointer value, Restore 表示 restore token_pointer value。

(1) 文法没有左递归。

```
Function P:boolean;
Begin
  Save;
  P:=true;
  If next_token=" begin" then
    If next_token=' d' then
      If next_token=' ;' then
        If X then
          If next_token=" end" then return;
        Restore;
        P:=false;
      End;
  Function X:boolean;
```

```
Begin
Save;
X:=true;
If next_token=' d' then
If next_token=' ;' then
If X then return;
Restore;
If next_token=' s' then
If Y then return;
Restore;
X:=false;
End;
Function Y:boolean;
Begin
Save;
Y=true;
If next_token=' ;' then
If next_token=' s' then
If Y then return;
Restore;
End;
```

(2) 消去文法左递归，并记为：

$P \rightarrow \text{begin } S \text{ end} \quad S \rightarrow A \mid CA \rightarrow V := EC \rightarrow \text{if } E \text{ then } S$

$E \rightarrow VE' \quad E' \rightarrow +VE' \quad | \quad \epsilon \quad V \rightarrow I$

Function P:boolean;

Begin

Save;

P:=true;

If next_token=" begin" then

If S then

If next_token=" end" then return; ;

Restore;

P:=false;

End;

Function A:boolean;

Beign

Save;

A:=true;

If V then

If next_token=" :" then

If E then return;

Restore;

A:=flase;

End;

Function S:boolean;

Beign

```
Save;  
S:=true;  
If A then return;  
Restore;  
If C then return;  
Restore;  
S:=false;  
End;  
Function C:boolean;  
Begin  
Save;  
C:=true;  
If next_token=" if" then  
If E then  
If next_token=" then" then  
If S then return;  
Restore;  
C:=false;  
End;  
Function E:boolean;  
Begin  
Save;  
E:=true;
```

```

If V then
If Ep then return;
Restore;
E:=false;
End;
Function Ep:boolean;
Being
Save;
Ep:=true;
If next_token='+' then
If V then
If E' then return;
Return;
End;

```

- 4. 解:

$\text{FIRST}_{\text{集}}$	$\text{FOLLOW}_{\text{集}}$
$\begin{array}{l} S \rightarrow aAB\# \\ \rightarrow bA\# \\ \rightarrow \epsilon \# \end{array}$	$\{a\} \#$ $\{b\} \#$ $\{\epsilon\} \#$
$\begin{array}{l} A \rightarrow aAb\# \\ \rightarrow \epsilon \# \end{array}$	$\{a\} \#$ $\{\epsilon\} \#$
$\begin{array}{l} B \rightarrow bB\# \\ \rightarrow \epsilon \# \end{array}$	$\{b\} \#$ $\{\epsilon\} \#$

- 5. 证: 因为是左递归文法, 所以必存在左递归的非终结符 A, 及形如 $A \rightarrow \alpha \mid \beta$ 的产生式, 且 $\alpha \in \text{FIRST}(\beta)$.

则 $\text{first}(\alpha) \cap \text{first}(\beta) \neq \emptyset$, 从而

$\text{first}(\alpha) \cap \text{first}(\beta) \neq \emptyset$, 即文法不满足 LL(1) 文法条件。得证。

- 6. 证: LL(1) 文法的分析句子过程的每一步, 永远只有唯一的分析动作可进行。现在, 假设 LL(1) 文法 G 是二义性文法, 则存在句子 a , 它有两个不同的语法树。即存在着句子 a 有两个不同的最左推导。从而可知, 用 LL(1) 方法进行句子 a 的分析过程中的某步中, 存在两种不同的产生式替换, 且均能正确进行语法分析, 即 LL(1) 分析动作存在不确定性。与 LL(1) 性质矛盾。所以, G 不是 LL(1) 文法。
- 7. 解:

(1) D 产生式两个候选式 f_D 和 f 的 first 集交集不为空, 所以不是 LL(1) 的。

(2) 此文法具有左递归性, 据第 5 题结论, 不是 LL(1) 的。

- 8. 解:

(1) 消除左递归性, 得:

$$S \rightarrow bZ11 \mid aZ21A \rightarrow bZ12 \mid aZ22Z11 \rightarrow bZ11 \mid \epsilon \quad Z12 \rightarrow bZ12$$

$$Z21 \rightarrow bZ11 \mid aZ21Z22 \rightarrow bZ12 \mid aZ22 \mid \epsilon$$

$$\text{消除无用产生式得: } S \rightarrow bZ11 \mid aZ21Z11 \rightarrow bZ11 \mid \epsilon \quad Z21 \rightarrow bZ11 \mid aZ21$$

此文法已满足 LL(1) 文法的三个条件,

所以 $G' [S]: S \rightarrow bZ11 \mid aZ21Z11 \rightarrow bZ11 \mid \epsilon \quad Z21 \rightarrow bZ11 \mid aZ21$

(2) G' 文法的各非终结符的 FIRST 集和 FOLLOW 集:

产生式	FIRST 集	FOLLOW 集
$S \rightarrow bZ11$	{b}	{#}
$\rightarrow aZ21$	{a}	
$Z11 \rightarrow bZ11$	{b}	{#}
$\rightarrow \epsilon$	{ ϵ }	
$Z21 \rightarrow bZ11$	{b}	{#}
$\rightarrow aZ21$	{a}	

LL(1) 分析表为:

	a	b	#
S	aZ21	bZ11	
Z11		bZ11	ϵ

Z21 aZ21 bZ11

- 9. 解:

(1)

产生式	first 集	follow 集
$S \rightarrow SaB$	{b}	{#, a, c}
$\rightarrow bB$	{b}	
$A \rightarrow S$	{b}	{c}
$\rightarrow a$	{a}	
$B \rightarrow Ac$	{a, b}	{#, a, c}

(2) 将 $S \rightarrow SaB \mid bB$ 改写为 $S \rightarrow bBS'$, $S' \rightarrow aBS' \mid \omega$, 可验证, 新文法是 LL(1) 的。

- 10. 解:

1) 为方便书写, 记: 〈布尔表达式〉为 A, 〈布尔因子〉为 B, 〈布尔二次量〉为 C, 〈布尔初等量〉为 D, 原文法可以简化为:

$A \rightarrow A \vee B \mid B \quad B \rightarrow B \wedge C \mid CC \rightarrow \neg D \mid DD \rightarrow (A) \mid \text{true} \mid \text{false}$,

显然, 文法含有左递归, 消去后等价 LL(1) 文法为:

$A \rightarrow BA' \quad A' \rightarrow \vee BA' \mid \omega \quad B \rightarrow CB' \quad B' \rightarrow \wedge CB' \mid \omega \quad C \rightarrow \neg D \mid DD \rightarrow (A) \mid \text{true} \mid \text{false}$

(2) 略

- 证: 若 LL(1) 文法 G 有形如 $B \rightarrow aAAb$ 的产生式, 且 $AT + \epsilon$ 及 AT^*ag , 根据 FIRST 集 FOLLOW 集的构造算法可知, FIRST(A) 中一切非 ϵ 加到 FOLLOW(A) 中, 则 $a \in \text{FOLLOW}(A)$; 又因为 $a \in \text{FIRST}(ag)$, 所以两集合相交非空, 因此, G 不是 LL(1) 文法; 与前提矛盾, 假设不成立, 得证。

- 解:

(1)

SA(a)b
 $S =$
 $A = < = <$
 $(== < < <$

```
<  
a>= <>>  
>  
)> >>>  
b >>
```

不是简单优先文法。

(2)

```
SRT() ∧ a,  
S > =  
R =  
T >  
( <= << < <  
) > >  
∧ > >  
a > >  
, < = < < <
```

是简单优先文法。

(3)

```
SR(a,)  
S = <<  
R >>  
( = <<  
a >>  
, = <<  
) >>
```

是简单优先文法。

◦ 首先消去无用产生式 $Z \rightarrow E$, $Z \rightarrow E+T$

```
SZT#i()  
S  
Z = =  
T > >  
# = < <<  
I > >  
( = < <<  
) > >
```

化简后的文法是简单优先文法；

- 解：

S	A	/	A
S		>	>
		=	
A	=	<	=
		<	
/		>	>
a		>	>

A 和/之间同时有关系=和<, 所以不是简单优先文法;

- 提示：分析教材中给出的算法，选择一种合适的表示给定文法的方法（尽量简单），使得对文法的输入比较简单同时（需要把输入转化为计算机语言表示，这种转化应该尽量简单），能够比较简单地构造3个基本关系矩阵（=， LEAD 和 LAST）。
 - 证明：设 $x_j x_{j+1} \dots x_i$ 是满足条件 $x_{j-1} < x_j = x_{j+1} = \dots = x_i > x_{i+1}$ 的最左子串。由=关系的定义，可知 $x_j x_{j+1} \dots x_i$ 必出现在某产生式的右部中。又因 $x_{j-1} < x_j$ 可知 x_{j-1} 与 x_j 不处于同一产生式，且 x_j 是某右部的首符。同理， x_i 为某产生式的尾符号。即存在产生式 $U \rightarrow x_j x_{j+1} \dots x_i$ 设 $ST * aUb$ 其中， $aT * \dots x_{j-1}$, $bT * x_{i+1} \dots$ 对于 aUb 可构造一语法树，并通过对 其剪枝（归约），直到 U 出现在句柄中。从而 $x_j x_{j+1} \dots x_i$ 必为句柄。反之，若 $x_j x_{j+1} \dots x_i$ 是句柄，由简单优先关系的定义，必满足上述条件。
 - 解：为描述方便，用符号表示各非终结符： $D = \langle \text{变量说明} \rangle$, $L = \langle \text{变量表} \rangle$, $V = \langle \text{变量} \rangle$, $T = \langle \text{类型} \rangle$, $a = \text{VAR}$, 则消去 V ，并采用分层法改写文法，得到：
$$D \rightarrow aW:T; W \rightarrow LL \rightarrow L, i \mid iT \rightarrow r \mid n \mid b \mid c$$

其全部简单优先关系是：

是简单优先文法。

- 证：设 $STna$, 我们对 n 用归纳法, 证明 a 不含两个非终结符相邻情况。 $n=1$ 时, STa , 即 $S \rightarrow a$ 是文法的产生式, 根据定义, 它不含上述情况。设 $n=k$ 时, 上述结论成立, 且设 $STkdAb$, 由归纳假设, A 两侧必为终结符。我们再进行一步推导, 得 $STkdAbTdb$, 其中, $A \rightarrow u$ 是文法中的产生式, 由定义, u 中不含两个非终结符相邻情况, 从而 db 两个非终结符相邻情况。得证。
- 证：由于 G 不是算符文法, G 中至少有一个产生式, 其右部含有两个非终结符相邻的情况。不失一般性, 设其形为 $U \rightarrow xABy$, $x, y \in V^*$, 由于文法不含无用产生式, 则必存在含有 U 的句型 dUb , 即存在推导 $ST \Rightarrow^* dUbTdxAByb$ 。得证。
- 文法为: $E \rightarrow E \uparrow A \mid AA \rightarrow A*T \mid A/T \mid TT \rightarrow T+V \mid T-V \mid VV \rightarrow i \mid (E)$
- 解:

(1) 构造算符优先矩阵:

-*	()	i	#
>>				
-	<	>	<>	
<<				
>				
*	>	<	>	<
<				
(<<	<	=	<	
)>>	>	>		
I>>	>	>		
#<<	<			

(2) 在 $(-, -)$ 、 $(-, *)$ 和 $(*, -)$ 处有多重定义元素, 不是算符优先文法;

(3) 改写方法:

- 将 $E \rightarrow E-T$ 中的减号与 $F \rightarrow P$ 中的赋值运算符强制规定优先关系;
- 或者将 $F-P$ 中的赋值运算符改为别的符号来表示;
- (1) 证明: 由设句型 $a = \dots Ua \dots$ 中含 a 的短语不含 U , 即存在 A , $A \Rightarrow^* ay$, 则 a 可归约为 $a = \dots Ua \dots \bar{u} \bar{y} \dots UA \dots = b$, b 是 G 的一个句型, 这与 G 是算符文法矛盾, 所以, a 中含有 a 的短语必含 U 。

(2) 的证明与 (1) 类似, 略。

- 证: (1) 对于 $a = \dots aU \dots$ 是句型, 必有 $ST \Rightarrow^* a (= \dots aU \dots) \mid T + \dots ab \dots$ 即在归约过程中, b 先于 a 被归约, 从而, $a < b$. 对于 (2) 的情况类似可以证明。
- 证明略。

- 证明略。
- 证明略。
- 证：(1) 用反证法。设没有短语包含 b 但是不包含 a ，则 a, b 一定同时位于某个短语中，从而必使得 a, b 同时位于同一产生式的右部，所以 $a=b$ ，与 G 是算符优先文法 (=与<不能并存) 矛盾。

(2)、(3) 类似可证。

- 证：只要证 u 中不含有除自身以外的素短语。设有这样的素短语存在，即存在 $bx \cdots by$ 是素短语，其中 $1 < x$ 或者 $y < n$ 之一成立。因素短语是短语，根据短语定义，则必有： $1 < xTbx-1 < bx$ 或 $y < nTby > by+1$ ，与 $bx-1=bx$ 及 $by=by+1$ 矛盾，得证。
- 提示：根据 27 题的结论，只要证 u 是句型 α 的短语，根据=关系的定义容易知道 u 是句型 α 的素短语。
- 证：与 28 题的不同点只是 a_0, a_{n+1} 可以是' #'，不影响结论。
- 证：设不能含有素短语，则只能是含有短语（不能含有终结符号），则该短语只能含有一个非终结符号，否则不符合算符文法定义，得证。
- 解：

(1) 算符优先矩阵：

/* ↑ () i#
+ ><< < > <>
* >>< < > <>
↑ >>< < > <>
(<<< < = <
) >>> > >
I >>> > >
<<< < <

(2) 用 Floyd 方法将优先矩阵线性化得到得的优先函数为：

/* ↑ () i#
F355 1771
G246 6161

- 解：用 Floyd 方法对已知的优先矩阵构造的优先函数为：

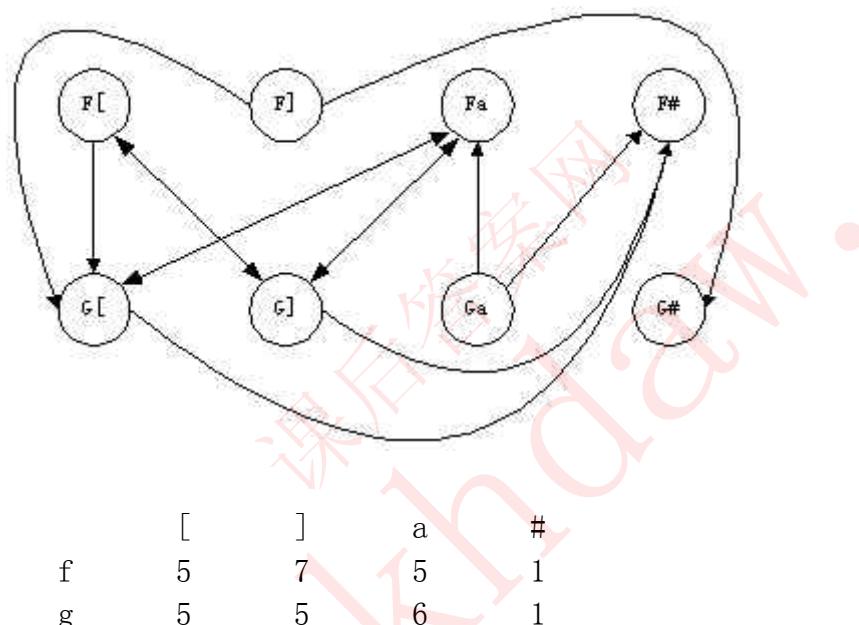
zbMLa()
f1567747
g1654667

- 解：

(1) 优先矩阵如下:

```
[] a#
[>=
]> >
<<
a <
>>
#<<<
```

(2) 用 Bell 方法求优先函数的过程如下:



(3) 显然, 文法不是算符优先文法, 所以不能线性化。

- 略。

35. 解:

(1) 识别全部活前缀的 DFA 如下: (以表格的形式来表示, 很容易可以转化为图的形式, 本章中其余的题目也是采用这种形式表示。)

状态 项目集 经过的符号到达的状态

S'	$\rightarrow \cdot S$	S	I_1
I_0	$S \rightarrow \cdot aSb$	a	I_2

$S \rightarrow \cdot aSc$

$S \rightarrow \cdot ab$

I1 $S' \rightarrow S \cdot$
 $S \rightarrow a \cdot Sb$

$S \rightarrow a \cdot Sc$

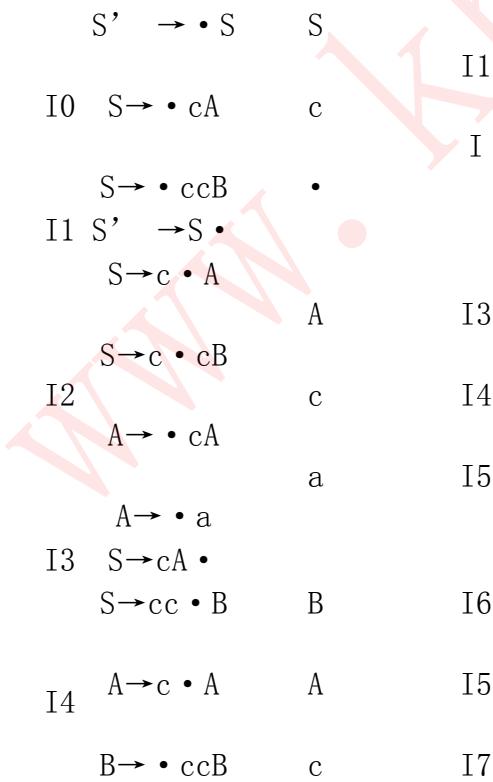
I2 $S \rightarrow a \cdot b$
 $S \rightarrow \cdot aSb$
 $S \rightarrow b$
 $S \rightarrow \cdot aSc$

I3 $S \rightarrow \cdot ab$
 $S \rightarrow aS \cdot b$
 $S \rightarrow b$
 $I5$

I4 $S \rightarrow ab \cdot$
 $I5 S \rightarrow aSb \cdot$
 $I6 S \rightarrow aSc \cdot$

(2) 识别全部活前缀的 DFA 如下:

状态 项目集 经过的符号到达的状态



	$B \rightarrow \cdot b$	b	I8
	$A \rightarrow \cdot cA$	a	I5
	$A \rightarrow \cdot a$		
I5	$A \rightarrow a \cdot$		
I6	$S \rightarrow ccB \cdot$		
	$B \rightarrow c \cdot cB$		
		C	I9
	$A \rightarrow c \cdot A$		
I7		A	I10
	$A \rightarrow \cdot cA$	a	I5
	$A \rightarrow \cdot a$		
I8	$B \rightarrow b \cdot$		
	$B \rightarrow cc \cdot B$		
		B	I11
	$A \rightarrow c \cdot A$		
	$B \rightarrow \cdot ccB$	A	I10
I9		c	I7
	$B \rightarrow \cdot b$	b	I8
	$A \rightarrow \cdot cA$	a	I5
	$A \rightarrow \cdot a$		
I10	$A \rightarrow cA \cdot$		
I11	$B \rightarrow ccB \cdot$		

所求的 LR(0) 项目规范族 $C = \{I0, I1, \dots, I11\}$

	(3)		
	状态	项目集	经过的符号到达的状态
	$S' \rightarrow \cdot S$	S	I1
I0	$S \rightarrow \cdot aSSb$	c	I2
	$S \rightarrow \cdot aSSS$	a	I3
	$S \rightarrow \cdot c$		
I1	$S' \rightarrow S \cdot$		
I2	$S \rightarrow c \cdot$		
I3	$S \rightarrow a \cdot SSb$	S	I4

$S \rightarrow a \cdot SSS$	c	I2
$S \rightarrow \cdot aSSb$	a	I3
$S \rightarrow \cdot aSSS$		
$S \rightarrow \cdot c$		
$S \rightarrow aS \cdot Sb$		
$S \rightarrow aS \cdot SS$	S	I5
I4 $S \rightarrow \cdot aSSb$	c	I2
$S \rightarrow \cdot aSSS$	a	I3
$S \rightarrow \cdot c$		
$S \rightarrow aSS \cdot b$		
$S \rightarrow aSS \cdot S$	S	I7
I5 $S \rightarrow \cdot aSSb$	a	I3
$S \rightarrow \cdot aSSS$	b	I6
$S \rightarrow \cdot c$	c	I2
I6 $S \rightarrow aSSb \cdot$		
I7 $S \rightarrow aSSS \cdot$		

(4)

状态 项目集 经过的符号到达的状态		
$S' \rightarrow \cdot S$	S	I1
I0 $S \rightarrow \cdot A$	A	I2
$A \rightarrow \cdot Ab$	a	I3
$A \rightarrow \cdot a$		
I1 $S' \rightarrow S \cdot$		
$S \rightarrow A \cdot$		
I2 $S \rightarrow A \cdot b$	b	I4
I3 $A \rightarrow a \cdot$		
I4 $S \rightarrow Ab \cdot$		

• 解：

(1) 是 LR(0) 文法，其 SLR(1) 分析表如下： $\text{FOLLOW}(S) = \{\#, b, c\}$

		ACTION				GOTO	
		a	b	c	#	S	
0	S2					1	
1				ACC			
2	S2	S4				3	
3		S5	S6				
4		R3	R3	R3			
5		R1	R1	R1			
6		R2	R2	R2			

(2) 是 LR(0) 文法，其 SLR(1) 分析表如下：

$\text{FOLLOW}(S) = \text{FOLLOW}(A) = \text{FOLLOW}(B) = \{\#\}$

				ACTION		GOTO		
				a	b	c	#	
0						S2	1	
1							ACC	
2	S5		S4					3
3		S5		S8		S7	R1	
4							R4	
5							R2	
6								
7		S5				S9		10
8			S8				R6	
9		S5		S8		S7		10 11
10							R3	
11							R5	

(3) 是 LR(0) 文法，其 SLR(1) 分析表如下： $\text{FOLLOW}(S) = \{\#, a, b, c\}$

		ACTION				GOTO	
		a	b	c	#	S	
0	S3			S2		1	
1				ACC			
2	R3	R3	R3	R3			
3	S3		S2			4	
4	S3		S2			5	

5	S3	S6	S2	7
6	R1	R1	R1	R1
7	R2	R2	R2	R2

(4) 因为 I_2 中含有冲突项目，所以不是 LR(0) 文法，其 SLR(1) 分析表如下：

$\text{FOLLOW}(S) = \{\#\} \cap \{b\} = \emptyset$ (所以可以用 SLR(1) 规则解决冲突)， $\text{FOLLOW}(A) = \{b, \#\}$

	ACTION			GOTO	
	a	b	#	S	A
0	S3			1	2
1			ACC		
2		S4	R1		
3		R3	R3		
4		R2	R2		

• 解：

状态 项目集 经过的符号到达的状态

	$S' \rightarrow \cdot S$	S	I1
I0	$S \rightarrow \cdot (SR)$	(I2
	$S \rightarrow \cdot a$	a	I3
I1	$S' \rightarrow S \cdot$		
	$S \rightarrow (\cdot SR)$	S	I4
I2	$S \rightarrow \cdot (SR)$	(I2
	$S \rightarrow \cdot a$	a	I3
I3	$S \rightarrow a \cdot$		
	$S \rightarrow (S \cdot R)$	a	I3
	$S \rightarrow \cdot (SR)$	(I2
I4	$S \rightarrow \cdot a$	R	I5
	$R \rightarrow \cdot, SR$,	I6
	$R \rightarrow \cdot)$)	I7
I5	$S \rightarrow (SR \cdot$		
I6	$R \rightarrow, \cdot SR$	S	I8

$S \rightarrow \bullet (SR \quad (\quad I2$
 $S \rightarrow \bullet a \quad a \quad I3$
I7 $R \rightarrow) \quad \bullet$
 $R \rightarrow, S \bullet R \quad) \quad I7$
I8 $R \rightarrow \bullet, SR \quad , \quad I6$
 $R \rightarrow \bullet) \quad R \quad I9$
I9 $R \rightarrow, SR \bullet$

LR(0) 分析表如下：

ACTION							GOTO	
	a	()	,	#	S	R	
0	S3	S2				1		
1					ACC		•	
2	S3	S2				4		
3	R2	R2	R2	R2	R2			
4	S3	S2	S7	S6			5	
5	R1	R1	R1	R1	R1			
6	S3	S2				8		
7	R4	R4	R4	R4	R4			
8			S7	S6			9	
9	R3	R3	R3	R3	R3			

可见是 LR(0) 文法。

• 解：

(1)

状态	项目集	经过的符号	到达的状态
	$S' \rightarrow \bullet S$		
I0	$S \rightarrow \bullet Sab$	S	I1
		b	I2
I1	$S' \rightarrow S \bullet$		acc
		a	
(冲突项目)	$S \rightarrow S \bullet ab$		I3

	$S \rightarrow b \cdot R$		
		R	I4
	$R \rightarrow \cdot S$		
		S	I5
I2	$R \rightarrow \cdot a$	a	I6
	$S \rightarrow \cdot S a b$	b	I2
	$S \rightarrow \cdot b R$		
I3	$S \rightarrow S a \cdot b$	b	I7
I4	$S \rightarrow b R \cdot$		r2
	$R \rightarrow S \cdot$		
I5		a	I3
	$S \rightarrow S \cdot a b$		
I6	$R \rightarrow a \cdot$		
I7	$S \rightarrow S a b \cdot$		

项目 I1, I5 同时具有移进和归约项目，对于 $I5=\{ R \rightarrow S \cdot, S \rightarrow S \cdot a b \}$, $\text{follow}(R)=\{a\}$, $\text{follow}(R) \cap \{a\}=\{a\}$, 所以 SLR(1) 规则不能解决冲突，从而该文法不是 SLR(1) 文法。

(2)

状态 项目集 经过的符号到达的状态

	$S' \rightarrow \cdot S$	S	I1
I0	$S \rightarrow \cdot a S A B$	a	I2
	$S \rightarrow \cdot B A$	B	I3
I1	$B \rightarrow \cdot b$	b	I4
	$S' \rightarrow S \cdot$		
	$S \rightarrow a \cdot S A B$	S	I5
I2	$S \rightarrow \cdot a S A B$	a	I2
	$S \rightarrow \cdot B A$	B	I3
	$B \rightarrow \cdot b$	b	I4
	$S \rightarrow B \cdot A$	A	I6
I3	$A \rightarrow \cdot a A$	a	I7
	$A \rightarrow \cdot B$	B	I8

	$B \rightarrow \cdot b$	b	I4
I4	$B \rightarrow b \cdot$		r5
	$S \rightarrow aS \cdot AB$	A	I9
	$A \rightarrow \cdot aA$	a	I7
I5	$A \rightarrow \cdot B$	B	I8
	$B \rightarrow \cdot b$	b	I4
I6	$S \rightarrow BA \cdot$		r2
	$A \rightarrow a \cdot A$	A	I10
	$A \rightarrow \cdot B$	B	I8
I7	$A \rightarrow \cdot aA$	a	I7
	$B \rightarrow \cdot b$	b	I4
I8	$A \rightarrow B \cdot$		r4
	$S \rightarrow aSA \cdot B$	B	I11
I9	$B \rightarrow \cdot b$	b	I4
I10	$A \rightarrow aA \cdot$		r3
I11	$S \rightarrow aSAB \cdot$		r1

不存在冲突项目，故该文法是 LR(0) 文法，也是 SLR(1) 文法。

SLR(1) 分析表如下：

	ACTION			#	GOTO		
	a	b	#		S	A	B
0	S2		S4		1		3
1				ACC			
2	S2	S4			5		3
3	S7	S4				6	8
4	R5	R5	R5				
5	S7	S4				9	8
6	R2	R2	R2				
7	S7	S4			10		8
8	R4	R4	R4				
9		S4					11
10	R3	R3	R3				
11	R1	R1	R1				

(3) 先求识别全部活前缀的 DFA:

状态	项目集 $S' \rightarrow \cdot S$	经过的符号	到达的状态
	$S \rightarrow \cdot aSb$	S	I1
I0	$S \rightarrow \cdot bSa$	a	I2
	$S \rightarrow \cdot ab$	b	I3
I1	$S' \rightarrow S \cdot$		acc
	$S \rightarrow a \cdot SB$		
	$S \rightarrow a \cdot b$	S	I4
I2	$S \rightarrow \cdot aSb$	b	I5
	$S \rightarrow \cdot bSa$	a	I2
	$S \rightarrow \cdot ab$		
	$S \rightarrow b \cdot Sa$		
	$S \rightarrow \cdot aSb$	S	I6
I3		a	I2
	$S \rightarrow \cdot bSa$	b	I3
	$S \rightarrow \cdot ab$		
I4	$S \rightarrow aS \cdot b$	b	I7
I5	$S \rightarrow ab \cdot$		r3
I6	$S \rightarrow bS \cdot a$	a	I8
I7	$S \rightarrow aSb \cdot$		r1
I8	$S \rightarrow bSa \cdot$		r2

不存在冲突项目，故该文法是 LR(0) 文法，也是 SLR(1) 文法。

SLR(1) 分析表如下：

	ACTION			GOTO
	a	b	#	
0	S2	S3		S
1			ACC	1
2	S2	S5		4
3	S2	S3		6
4		S7		
5	R3	R3	R3	
6	S8			

7	R1	R1	R1
8	R2	R2	R2

(4) 先求识别全部活前缀的 DFA:

状态 项目集 经过的符号到达的状态

	$S' \rightarrow \cdot S$	S	$I1$
I0	$S \rightarrow \cdot aA$	a	$I2$
	$S \rightarrow \cdot bB$	b	$I3$
I1	$S' \rightarrow S \cdot$		acc
	$S \rightarrow a \cdot A$		$I4$
I2		A	
	$A \rightarrow \cdot cAd$		$I5$
(冲突)		c	
	$A \rightarrow \cdot$		$r4$
	$S \rightarrow b \cdot B$		$I6$
I3		B	
	$B \rightarrow \cdot cBdd$		$I7$
(冲突)		c	
	$B \rightarrow \cdot$		$r6$
I4	$S \rightarrow aA \cdot$		$r1$
	$A \rightarrow c \cdot Ad$		$I8$
I5		A	
	$A \rightarrow \cdot cAd$		$I5$
(冲突)		c	
	$A \rightarrow \cdot$		$r4$
I6	$S \rightarrow bB \cdot$		$r2$
	$B \rightarrow c \cdot Bdd$		$I9$
I7		B	
	$B \rightarrow c \cdot Bdd$		$I7$
(冲突)		c	
	$B \rightarrow \cdot$		$r6$
I8	$A \rightarrow cA \cdot d$	d	$I10$
I9	$B \rightarrow cB \cdot dd$	d	$I11$
I10	$A \rightarrow cAd \cdot$		$r3$
I11	$B \rightarrow cBd \cdot d$	d	$I12$
I12	$B \rightarrow cBdd \cdot$		$r5$

因为 $\text{follow}(A) = \text{follow}(B) = \{\#, d\}$, 所以冲突项目 $I2, I3, I5, I7$ 可以用 SLR(1) 规则得以解决, 从而该文法为 SLR(1) 文法。其 SLR(1) 分析表如下:

ACTION	GOTO
--------	------

a	b	c	d	#	SAB
0	S2	S3		1	
1				Acc	
2	S5	R4	R4	4	
3	S7	R6	R6	6	
4				R1	
5	S5	R4	R4	8	
6				R2	
7	S7	R6	R6	9	
8				S10	
9				S11	
10				R3	R3
11				S12	
12				R5	R5

(5) 解: 原文法等价化为 $q_1 \rightarrow q_2$, $q_1 \rightarrow q_3$, $q_2 \rightarrow q_4; q_5$, $q_4 \rightarrow \text{begin D}$, $q_4 \rightarrow q_4; D$, $q_5 \rightarrow S \text{ end}$, $q_5 \rightarrow S; q_5$, $q_3 \rightarrow \text{begin} q_5$,

先求识别全部活前缀的 DFA:

状态	项目集	经过的符号到达的状态
	$q_1' \rightarrow \cdot q_1$	
	$q_1 \rightarrow \cdot q_2$	q_1
	$q_1 \rightarrow \cdot q_3$	q_2
I0	$q_2 \rightarrow \cdot q_4; q_5$	q_3
	$q_3 \rightarrow \cdot \text{begin} q_5$	q_4
	$q_4 \rightarrow \cdot \text{begin D}$	begin
	$q_4 \rightarrow \cdot q_4; D$	
I1	$q_1' \rightarrow q_1 \cdot$	ACC
I2	$q_1 \rightarrow q_2 \cdot$	R1
I3	$q_1 \rightarrow q_3 \cdot$	R2
	$q_2 \rightarrow q_4 \cdot; q_5$;
I4	$q_4 \rightarrow q_4 \cdot; D$	I6
	$q_3 \rightarrow \text{begin} \cdot q_5$	q5
I5	$q_3 \rightarrow \text{begin} \cdot D$	D
		I8

	q5 → • Send	S	I9
	q5 → • S; q5		
	q2 → q4; • q5	q5	I10
	q4 → q4; • D		
I6		D	I11
	q5 → • Send	S	I9
	q5 → • S; q5		
I7	q3 → begin q5 •		R8
I8	q3 → begin D •		R4
	q5 → S • end	end	I12
I9			
	q5 → S • ; q5	;	I13
I10	q2 → q4; q5 •		R3
I11	q4 → q4; D •		R5
I12	q5 → Send •		R6
	q5 → s; • q5		
I13	q5 → • Send	q5	I14
		S	I9
	q5 → • S; q5		
I14	q5 → s; q5 •		R7

不存在冲突项目，故该文法是 LR(0) 文法，也是 SLR(1) 文法。

SLR(1) 分析表如下：

	ACTION	GOTO
	; begin D S end #	q1q2q3q4q5
0		1 2 3 4
1	Acc	
2	R1	
3	R2	
4 S6		
5	S8S9	7
6	S7S9	10
7		R8
8		R4
9 S13	S12	
10		R3
11 R5		
12		R6

13

S9

14

14

R7

(6) 解: 原文法可化为等价形式: $q_1 \rightarrow \text{begin } q_2 \ q_3 \ \text{end}$, $q_2 \rightarrow q_2 \ d;$,
 $q_2 \rightarrow \epsilon$, $q_3 \rightarrow q_3; q_4$, $q_3 \rightarrow q_4$, $q_4 \rightarrow S$, $q_4 \rightarrow \epsilon$,

先求识别全部活前缀的 DFA:

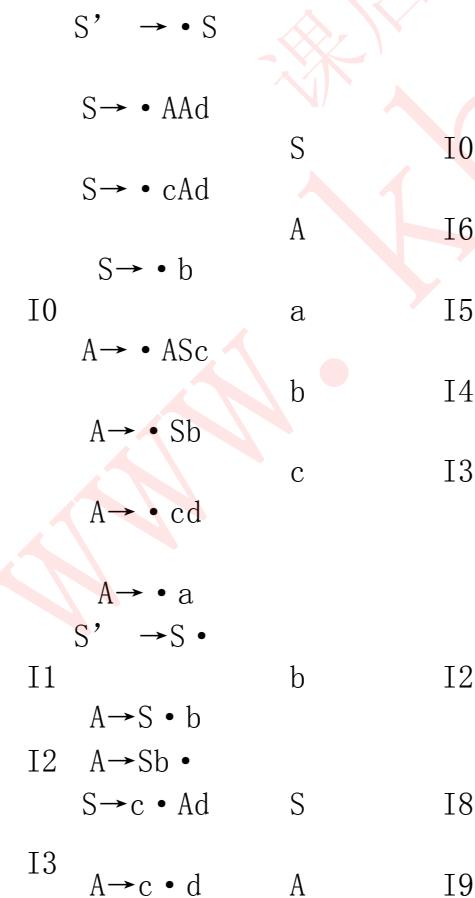
状态	项目集	经过的符号到达的状态
I0	$q_1' \rightarrow \cdot q_1$	q_1
I1	$q_1 \rightarrow \cdot \text{begin} q_2 q_3 \text{end}$	begin
I2	$q_2 \rightarrow \cdot q_2 d;$	q_2
I3	$q_3 \rightarrow \cdot q_3; q_4$	q_3
(冲突项目)	$q_3 \rightarrow \cdot q_4$	q_4
I4	$q_4 \rightarrow \cdot S$	S
I5	$q_1 \rightarrow \text{begin} q_2 q_3 \cdot \text{end}$	end
I6	$q_3 \rightarrow q_3 \cdot; q_4$	$;$
I7	$q_3 \rightarrow q_4 \cdot$	$R5$
I8	$q_4 \rightarrow S \cdot$	$R6$
(冲突项目)	$q_4 \rightarrow \cdot$	$R1$
I9	$q_3 \rightarrow q_3; q_4 \cdot$	$I9$
I10	$q_2 \rightarrow q_2 d \cdot;$	$;$
I11	$q_2 \rightarrow q_2 d; \cdot$	$R2$

因为 $\text{follow}(q_4) = \{\text{end}, \#\}$ ，故冲突项目可以通过 SLR(1) 规则来解决，从而文法为 SLR(1) 文法。SLR(1) 分析表如下：

	action	goto
	beginend d ; S # q1q2q3q4	
0 S2		1
1		
2	R3 R3 R3R3	3
3	R7 S10R7S6	4 5
4	S7 S8	
5	R5 R5	
6	R6 R6	
7		R1
8	R7 R7S6	9
9	R4 R4	
10		
11	R2 R2 R2R2	

39. 解：识别活前缀的 DFA 及 LR(0) 分析表：

状态 项目集 经过的符号到达的状态



A \rightarrow • ASc a I5

A \rightarrow • Sb b I4

A \rightarrow • cd c I3

A \rightarrow • a d I7

S \rightarrow • AAd

S \rightarrow • cAd

S \rightarrow • b
I4 S \rightarrow b •
I5 A \rightarrow a •
S \rightarrow A • Ad

A \rightarrow A • Sc

A \rightarrow • ASc S I11

A \rightarrow • Sb A I10
I6 A \rightarrow • cd a I5
A \rightarrow • a b I4
S \rightarrow • AAd c I3

S \rightarrow • cAd

S \rightarrow • b
I7 A \rightarrow cd •
I8 A \rightarrow S • b b I2
S \rightarrow cA • d S I11

A \rightarrow A • Sc A I10

I9 S \rightarrow A • Ad a I5

S \rightarrow • AAd b I4

S \rightarrow • cAd c I3

d I14

$S \rightarrow \bullet b$
 $A \rightarrow \bullet ASc$
 $A \rightarrow \bullet Sb$
 $A \rightarrow \bullet cd$
 $A \rightarrow \bullet a$
 $S \rightarrow AA \bullet d$
 $A \rightarrow A \bullet Sc$
 $S \rightarrow A \bullet Ad$ S I11
 $S \rightarrow \bullet AAd$ A I10
 $S \rightarrow \bullet cAd$ a I5
I10 $S \rightarrow \bullet b$ b I4
 $A \rightarrow \bullet ASc$ c I3
 $A \rightarrow \bullet Sb$ d I13
 $A \rightarrow \bullet cd$
 $A \rightarrow \bullet a$
 $A \rightarrow AS \bullet c$ b I2
I11 $A \rightarrow S \bullet b$ c I12
I12 $A \rightarrow ASc \bullet$
I13 $S \rightarrow AAd \bullet$
I14 $S \rightarrow cAd \bullet$
ACTION GOTO
a b c d # S A
0 s5s4 s3 1 6
1 s2 acc
2 r5r5 r5 r5 r5
3 s5s4 s3 s7 8 9
4 r3r3 r3 r3 r3
5 r7r7 r7 r7 r7
6 s5s4 s3
7 r6r6 r6 r6 r6 1110

```

8 s2
9 s5s4 s3 s14    1110
10s5s4 s3 s13   1110
11 s2s12
12r4r4 r4 r4 r4
13r1r1 r1 r1 r1
14r2r2 r2 r2 r2

```

SLR(1) 分析法: FOLLOW(S)={c, b} FOLLOW(A)={a, b, c, d}

ACTION	GOTO
a b c d # S A	
0 s5s4 s3	1 6
1 s2	acc
2 r5r5 r5 r5	
3 s5s4 s3 s7	8 9
4 r3 r3	
5 r7r7 r7 r7	
6 s5s4 s3	
7 r6r6 r6 r6	1110
8 s2	
9 s5s4 s3 s14	1110
10s5s4 s3 s13	1110
11 s2s12	
12R4r4 r4 r4	
13 r1 r1	
14 r2 r2	

两表的异同: 两表都用 ACTION 表和 GOTO 表反映了移进、归约 (包括接受) 的状态和动作。不同之处在于 SLR(1) 增加了在归约的时候考虑向前符号 a 以解释可能出现的“移进——归约”冲突; SLR(1) 的分析较稀疏些, 原因是填写归约项时, 并不是在一状态对应行上全部填写归约动作, 而是考虑了相应非终结符的 FOLLOW 集因素。另外, 在分析效率上 SLR(1) 分析的效率更高一些, 因为在发现错误方面, SLR(1) 比 LR(0) 分析发现的更早些。

40. 解: 求 LR(1) 项目集和状态转换表:

状态	项目集	经过的符号到达的状态
S'	$\rightarrow \cdot S \#$	S
I0	$S \rightarrow \cdot A \#$	A
		I1
		I2

	A \rightarrow • BA#	B	I3
	A \rightarrow • #	a	I4
	B \rightarrow • aB#, a, b	b	I5
	B \rightarrow • b#, a, b		
I1	S' \rightarrow S • #		
I2	S \rightarrow A • #		
	A \rightarrow B • A#		
		A	I6
	A \rightarrow • #	B	I3
I3	A \rightarrow • BA#	a	I4
	B \rightarrow • aB#, a, b	b	I5
	B \rightarrow • b#, a, b		
	B \rightarrow a • B#, a, b	B	I7
I4	B \rightarrow • aB#, a, b	a	I4
	B \rightarrow • b#, a, b	b	I5
I5	B \rightarrow b • #, a, b		
I6	A \rightarrow BA • #		
I7	B \rightarrow aB • #, a, b		

相应的 LR(1) 分析表为：

STATE	ACTION			GOTO		
	a	b	#	S	A	B
0	S4	S5	R3	1	2	3
1			ACC			
2			R2			
3	S4	S5	R3	6	3	
4	S4	S5			7	
5	R5	R5	R5			
6			R2			
7	R4	R4	R4			

用 LR(1) 分析表对输入符号串 abab 的分析过程：

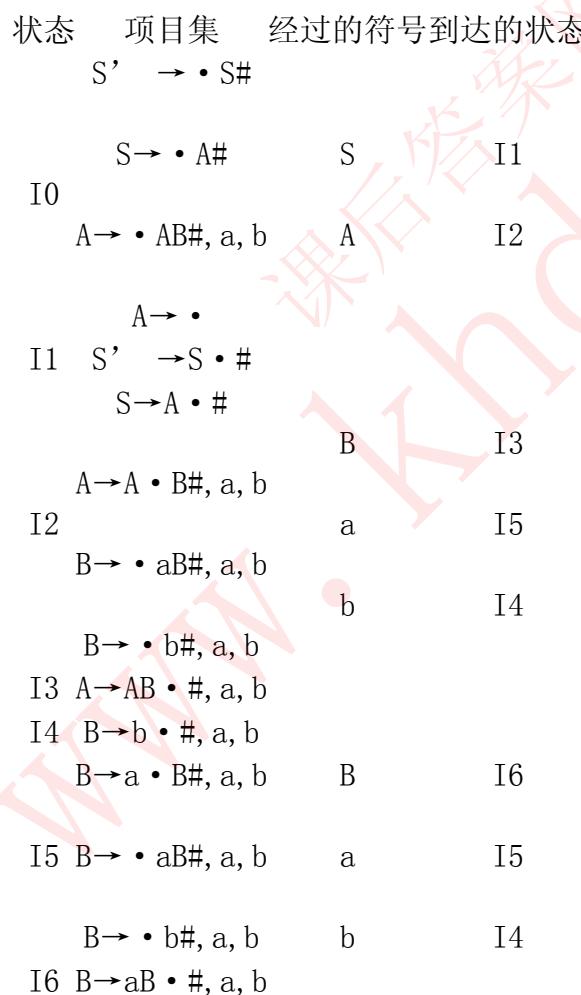
步骤状态栈中符号余留符号分析动作下一状态

1 0 # abab# S4 4

2	04	#a	bab#	S5	5
3	045	#ab	ab#	R5	7
4	047	#aB	ab#	R4	3
5	03	#B	ab#	S4	4
6	034	#Ba	b#	S5	5
7	0345	#Bab	#	R4	7
8	0347	#BaB	#	R4	3
9	033	#BB	#	R3	6
10	0336	#BBA	#	R2	6
11	036	#BA	#	R2	1
12	01	#A	#	acc	

41. 解：

(1) 求 LR(1) 项目集和状态转换图：



相应的 LR(1) 分析表为：

STATE	ACTION	GOTO
-------	--------	------

	a	b	#	S	A	B
0	R3	R3	R3	1	2	
1			Acc			
2	S5	S4	R1			
3	R2	R2	R2			
4	R5	R5	R5			
5	S5	S4				6
6	R4	R4	R4			

表中没有多从定义的元素，所以文法是 LR(1) 文法。

(2) LR(1) 分析法：

状态 项目集 经过的符号到达的状态

	$E' \rightarrow \cdot E\#$		
		E	I1
	$E \rightarrow \cdot E + T \# / +$		
		T	I1
I0	$E \rightarrow \cdot T \# / +$	(I5
		a	I4
	$T \rightarrow \cdot (E) \# / +$		
	$T \rightarrow \cdot a \# / +$		
	$E' \rightarrow E \cdot \#$	+	I2
I1			
	$E \rightarrow E \cdot + T \# / +$	T	I3
	$E \rightarrow E + \cdot T \# / +$		
I2	$T \rightarrow \cdot (E) \# / +$	(I5
		a	I4
	$T \rightarrow \cdot a \# / +$		
I3	$E \rightarrow E + T \cdot \# / T$		
I4	$T \rightarrow a \cdot \# / +$		
	$T \rightarrow (\cdot E) \# / +$	E	I7
	$E \rightarrow \cdot E + T + /)$	C	I8
		T	I9
I5	$E \rightarrow \cdot T) / +$	a	I12
	$T \rightarrow \cdot (E) + /)$		
	$T \rightarrow \cdot a + /)$		
I6	$E \rightarrow T \cdot \# / +$		

	$T \rightarrow (E \bullet) \# / +)$		I10
I7	$E \rightarrow E \bullet + T / +)$	+	I11
	$T \rightarrow (\bullet E) / +)$	(I8
	$E \rightarrow \bullet E + T / +)$	a	I12
I8	$E \rightarrow \bullet T) / +)$	E	I14
	$T \rightarrow \bullet (E) / +)$	T	I9
	$T \rightarrow \bullet a / +)$		
I9	$E \rightarrow T \bullet / +)$		
I10	$T \rightarrow (E) \bullet \# / +)$	(I8
	$E \rightarrow E + \bullet T / +)$		
I11	$T \rightarrow \bullet (E) / +)$	T	I13
	$T \rightarrow \bullet a / +)$	a	I12
I12	$T \rightarrow a \bullet / +)$		
I13	$E \rightarrow E + T \bullet / +)$		
	$T \rightarrow (E \bullet) / +)$	+	I11
I14	$E \rightarrow E \bullet + T / +)$		I15
	$I15 T \rightarrow (E) \bullet / +)$		

LALR(1) 分析: (合并同心集)

状态	项目集	经过的符号到达的状态
	$E' \rightarrow \bullet E \#$	
I0	$E \rightarrow \bullet E + T \# / +)$	E
	$E \rightarrow \bullet T \# / +)$	I1
	$T \rightarrow \bullet (E) \# / +)$	a
	$T \rightarrow \bullet a \# / +)$	I4/I12
I1	$E' \rightarrow E \bullet \#$	
	$E \rightarrow E \bullet + T \# / +)$	+
	$E \rightarrow E + \bullet T \# / +)$	I2/I11
	$E \rightarrow E \bullet + T \# / +)$	T
	$E \rightarrow E + \bullet T \# / +)$	I3/I13
I2/I11	$T \rightarrow \bullet (E) \# / + /)$	(
		I5/I8

	$T \rightarrow \bullet a \# / + /)$	a	I4/I12
I3/I13	$E \rightarrow E + T \bullet /) / + / \#$		
I4/I12	$T \rightarrow a \bullet / + /) / \#$		
	$T \rightarrow (\bullet E) \# / + /)$	T	I6/I9
	$E \rightarrow \bullet E + T + /)$	(I5/I8
I5/I8	$E \rightarrow \bullet T) / +$	E	I7/I14
	$T \rightarrow \bullet (E) + /)$	a	I4/I12
	$T \rightarrow \bullet a + /)$		
I6/I9	$E \rightarrow T \bullet + /) / \#$		
	$T \rightarrow (E \bullet) + /) / \#$	+	I2/I11
I7/I14	$E \rightarrow E \bullet + T + /)$)	I10/I15
I10/I15	$T \rightarrow (E \bullet + /) / \#$		
LR(1)	ACTION	GOT	
)	+ () a # E T	0	
0	S5 S4 1 6		
1	S2 ACC		
2	S5 S4 3		
3	R1 R1		
4	R4 R4		
5	S8 S12 7 9		
6	R2 R2		
7	S11 S10		
8	S8 S12 14 9		
9	R2 R2		
10	R3 R3		
11	S8 S12 13		
12	R4 R4		
13	R1 R1		
14	S11 S15		
15	R3 R3		

可以看出，表中无冲突项，所以是 LR(1) 文法；LALR(1) 分析表：

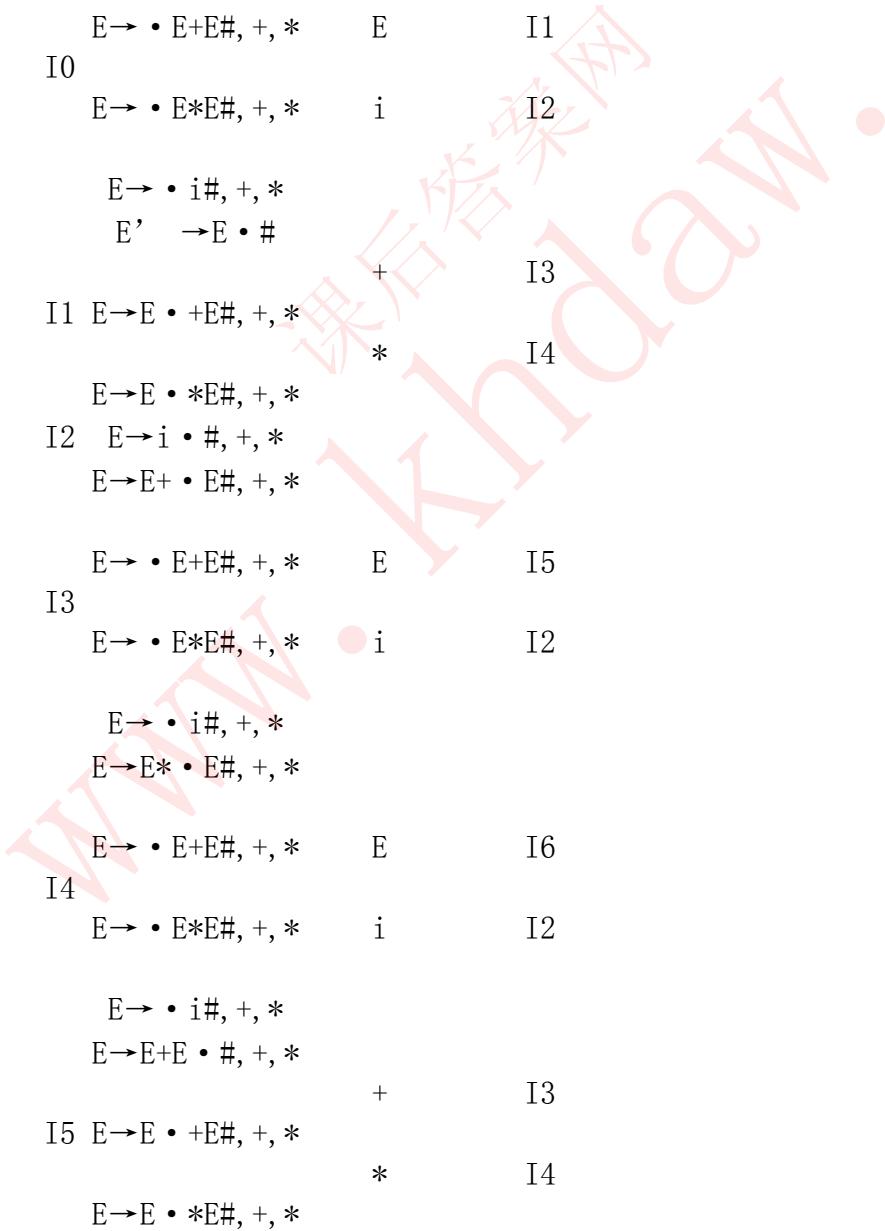
LR(1)		ACTION		GOTO			
)	+	())	a	#	E	T
0		S5/S8		S4/S12		1	6/9
1	S2/S11			ACC			

2/11	S5/S8	S4/S12	3/13
3/13	R1	R1	R1
4/12	R4	R4	R4
5/8	S5/S8	S4/S12	7/14 6/9
6/9	R2	R2	R2
10/15	R3	R3	R3
7/14	S2/S11	S10/S15	

• 解：

(1) 求 LR(1) 项目集和状态转换图：

状态 项目集 经过的符号到达的状态
 $E' \rightarrow \cdot E\#$



$E \rightarrow E * E \bullet \#, +, *$		
	+	I3
I6 $E \rightarrow E \bullet + E \#, +, *$		
	*	I4

$E \rightarrow E \bullet * E \#, +, *$

依据以上图求出该文法的 LR(1) 分析表知道由于项目 I5, I6 导致了有多重定义的元素, 所以不是 LR(1) 文法。事实上, 从文法本身可以看出它是二义性的, 因此不可能是 LR(1) 文法。等价的 LR(1) 文法为: $E \rightarrow E + T \mid T \rightarrow T * F \mid F \rightarrow i$ 。另外, 对原文法的冲突项来说, 若考虑算术运算符的运算优先级别, 以及结合方式, 上述冲突是可解决的。例如, 假设表达式运算满足左结合律(即 $a+b+c=(a+b)+c$ 而不是右结合律: $a+b+c=a+(b+c)$), 及*运算和优化优先级高于+运算, 上述文法相应的 LR(1) 分析表为

状态	ACTION	GOT
		0
i + * #	E	
0 s2		1
1 s3s4acc		
2 r3r3 r3		
3 s2		5
4 s2		6
5 r1s4 r1		
6 r2r2 r2		

(2) 解: LR(1):

状态 项目集 经过的符号到达的状态

$S' \rightarrow \bullet S \#$		
$S \rightarrow \bullet a S a \#$	S	I1
I0 $S \rightarrow \bullet b S b \#$	a	I2
$S \rightarrow \bullet a \#$	b	I3
$S \rightarrow \bullet b \#$		
I1 $S' \rightarrow S \bullet \#$		
$S \rightarrow a \bullet S a \#$	S	I4
I2 $S \rightarrow a \bullet \#$	a	I7
$S \rightarrow \bullet a S a a$	b	I6

$S \rightarrow \bullet bSba$

$S \rightarrow \bullet aa$

$S \rightarrow \bullet ba$

$S \rightarrow b \bullet Sb\#$

$S \rightarrow b \bullet \#$

	S	I11
I3	a	I14
	b	I13
	$S \rightarrow \bullet bSbb$	

	S	I15
I4	a	I15
I5	$S \rightarrow aSa \bullet \#$	
	$S \rightarrow b \bullet Sba$	

	S	I10
I6	a	I14
	b	I13
	$S \rightarrow \bullet ab$	

	S	I10
	a	I14
	b	I13
	$S \rightarrow \bullet bb$	
	$S \rightarrow a \bullet Saa$	

	S	I8
I7	a	I7
	b	I6
	$S \rightarrow \bullet aa$	

	S	I8
I8	a	I9
I9	$S \rightarrow aSa \bullet a$	
I10	b	I15
I11	b	I12
I12	$S \rightarrow bSb \bullet \#$	

$S \rightarrow b \cdot Sbb$

	$S \rightarrow b \cdot b$		
	$S \rightarrow \cdot aSab$	S	I16
I13		a	I14
	$S \rightarrow \cdot bSbb$	b	I13
	$S \rightarrow \cdot ab$		
	$S \rightarrow \cdot bb$		
	$S \rightarrow a \cdot Sab$		
	$S \rightarrow a \cdot b$		
	$S \rightarrow \cdot aSaa$	S	I18
I14		a	I7
	$S \rightarrow \cdot bSba$	b	I6
	$S \rightarrow \cdot aa$		
	$S \rightarrow \cdot ba$		
I15	$S \rightarrow bSb \cdot a$		
I16	$S \rightarrow bS \cdot bb$	b	I17
I17	$S \rightarrow bSb \cdot b$		
I18	$S \rightarrow aS \cdot ab$	a	I19
I19	$S \rightarrow aSa \cdot b$		

有移进——归约冲突，不是 LR(1) 文法。

(3) 求 LR(1) 项目集和状态转换图：

状态	项目集	经过的符号到达的状态
	$S' \rightarrow \cdot S\#$	
	$S \rightarrow \cdot V := E\#$	S
		I1
I0	$S \rightarrow \cdot LS\#$	V
		I2
	$L \rightarrow \cdot I : I$	L
		I3
	$V \rightarrow \cdot I :$	I
I1	$S' \rightarrow S \cdot \#$	I4
I2	$S \rightarrow V \cdot := E\#$:
		I5

S→L • S#			
	S→ • V:=E#	S	I6
I3	S→ • LS#	L	I3
	V→ • I:	I	I4
	L→ • I:I		
	L→ I • :I		
I4	:		I7
	V→ I • :		
I5	S→ V: • =E#	=	I8
I6	S→ LS • #		
I7	L→ I: • I		
I8	S→ V:= • E#	E	I9
I9	S→ V:=E • #		

依据以上图求出该文法的 LR(1) 分析表知道由于项目 I4 导致了有多重定义的元素，所以不是 LR(1) 文法。

(4) 略。

- 证：根据第 3 章“词法分析及词法分析程序”，我们知道任一正规集可以由某一正规式 r 表示，而对于任一正规式 r ，都可以构造一个 DFA，并且两者在表述语言的意义下等价。根据这个 DFA，我们可以很容易地构造一个分析表。方法是对于 DFA 中的每一个子图，如果从状态 I_i ，经过了终结符号 a ，到达状态 I_j ，则在分析表中有 $ACTION(I_i, a) = S_j$ ，如果 S_j 是一个终结状态，则置 $ACTION(I_i, a) = ACC$ 。
- 解：SLR(1) 分析表比 LR(0) 分析表在采用一定形式存储时（如稀疏矩阵），会少一些存储空间。另外，在分析进行到归约动作时，LR(0) 无论下一符号是什么（即使是错误的输入），都将先进行归约，然后才判断下一输入符是否正确；而 SLR(1) 在进行归约时还要先看看下一符号是否正确，否则将报错。从而可知，对于有错误的输入串，SLR(1) 分析效率要高一些。
- 证：考察 LR 分析器的总控程序，可以知道访问 GOTO 分析表元素的可能只有两种，下面分情况讨论：
- 先访问 $ACTION(S_m, a_i) = " 移进 "$ ，再访问 $GOTO(S_m, a_i) = S_{m+1}$ ，其中 a_i 是一个终结符。查看构造 LR 分析表的算法的条目（1），可以知道当 a_i 是一个终结符时，填写 $ACTION(S_m, a_i) = " 移进 "$ 总是伴随着填写 $GOTO(S_m, a_i) = S_{m+1}$ ，所以这种情况下结论成立。
- 先访问 $ACTION(S_m, a_i) = r_j$ ，其中 a_i 是一个终结符， r_j 意指按文法的第 j 个产生式 $A \rightarrow X_m - r + 1 X_m - r + 2 \dots X_m$ 进行归约，再访问 $GOTO(S_m - r, A) = S_1$ ，

其中 A 是一个非终结符。查看构造 LR 分析表的算法的条目 (1) , 可以知道当 a_i 是一个非终结符时, 填写 $GOTO(S_m, a_i) = S_{m+1}$, 即只有归约出了 a_i 时, 再状态转换到 S_{m+1} , 和总控程序情况一致, 所以这种情况下结论成立。

综上所述, 结论成立。

46. 证明: 在文法 $G[S]$ 中, 只有非终结符 S 有两个候选式, $AaAb$ 及 $BbBa$, 而两个候选的 FIRST 集分别为 $\{a\}$ 和 $\{b\}$, 它们不相交, 所以, 文法 $G[S]$ 满足 LL (1) 文法的条件, 是 LL (1) 文法。在 SLR (1) 方法识别活前缀的 DFA 中, 项目 $I_0 = \{S' \rightarrow \cdot S, S \rightarrow \cdot AaAb, S \rightarrow \cdot BbBa, A \rightarrow e \cdot, B \rightarrow e \cdot\}$ 中出现归约-归约冲突, 而 $follow(A) = follow(B) = \{a, b\}$, 用 SLR (1) 方法不能解决冲突。

47. 略。

48. 略。

第五章 习题解答

5.1 解：

- 属性文法是文法符号带有语义属性的前后文无关文法；
- 属性翻译文法首先是对文法的属性依赖关系作出限制，不允许出现属性的直接或间接的循环定义，即要求属性文法是良定义的；其次还应将属性定义规则改造为计算属性值的语义程序，即将静态的定义规则改写为可动态执行的语义动作；
- 属性文法是静态描述，而属性翻译文法是动态描述，有语义动作。

5.2 解：

- 综合属性有： $Z. aZ. pX. dX. c$

继承属性有： $X. b$

- 属性依赖出现了循环；

5.3 解：S 属性文法一定是 L 属性文法，因为前者是在后者基础上加上限制条件，即非终结符只有综合属性；反之当然不正确。

5.4 解：

- $A-BC+*DE-\hat{}$
- $ad*c+d/e+f*g+$
- $ax+4 \leq cd3*/\backslash\backslash$
- $de*c+b/\backslash a\backslash/f\hat{}$
- $s0=; i1=; i100 \leq BZ s s ii*+=; ii1+=; BR S2'$

5.5 解：

- $a+b*c$
- $a*(b-c)-(c+d)/e$
- 有误
- $if(a < b) x = (a-b)^c; else g = h;$

5.6 略

5.7 略

5.8 解：

• $(+, B, C, T1)$

$(*, A, T1, T2)$

$(+, T2, D, T3)$

$(=, T3, 0, X)$

2) 如下所示:

- 当前句型 (方框括起来部分为句柄) $A/\backslash(BV(CVD/\backslash?F))$ 用产生式 $Expr \rightarrow iden$ 归约, 得

(1) $Expr. TC \rightarrow (jnz, A, 0, 0);$

(2) $Expr. FC \rightarrow (j, 0, 0, 0);$

- 当前句型 $Expr/\backslash(BV(CVD/\backslash?F))$ 用产生式 $Expr^{\wedge} \rightarrow Expr' / \backslash'$ 归约, 得

(1) $(jnz, A, 0, 3);$

(2) $Expr^{\wedge}. FC \rightarrow (j, 0, 0, 0);$

- 当前句型 $Expr^{\wedge}(BV(CVD/\backslash?F))$ 用产生式 $Expr \rightarrow iden$ 归约, 得

(1) $(jnz, A, 0, 0);$

(2) $Expr^{\wedge}. FC \rightarrow (j, 0, 0, 0);$

(3) $Expr. TC \rightarrow (jnz, B, 0, 0);$

(4) $Expr. FC \rightarrow (j, 0, 0, 0);$

- 当前句型 $Expr^{\wedge}(ExprV(CVD/\backslash?F))$ 用产生式 $ExprV \rightarrow Expr' V'$ 归约, 得

(1) $(jnz, A, 0, 3);$

(2) $Expr^{\wedge}. FC \rightarrow (j, 0, 0, 0);$

(3) $ExprV. TC \rightarrow (jnz, A, 0, 0);$

(4) $(j, 0, 0, 5);$

- 当前句型 $Expr^{\wedge}(ExprV(CVD/\backslash?F))$ 用产生式 $Expr \rightarrow iden$ 归约, 得

(1) $(jnz, A, 0, 3);$

(2) $Expr^{\wedge}. FC \rightarrow (j, 0, 0, 0);$

(3) Exprv. TC \rightarrow (jnz, A, 0, 0) ;

(4) (j, 0, 0, 5) ;

(5) Expr. TC \rightarrow (jnz, C, 0, 0) ;

(6) Expr. FC \rightarrow (j, 0, 0, 0) ;

- 当前句型 $\text{Expr}^{\wedge} (\text{Exprv} (\text{ExprvD}/\text{?F}))$, 用产生式 $\text{Exprv} \rightarrow \text{Expr}' V'$ 归约, 得

(1) (jnz, A, 0, 3) ;

(2) Expr $^{\wedge}$. FC \rightarrow (j, 0, 0, 0) ;

(3) Exprv. TC \rightarrow (jnz, A, 0, 0) ;

(4) (j, 0, 0, 5) ;

(5) Exprv. TC \rightarrow (jnz, C, 0, 0) ;

(6) (j, 0, 0, 7) ;

- 当前句型 $\text{Expr}^{\wedge} (\text{Exprv} (\text{ExprvD}/\text{?F}))$, 用产生式 $\text{Expr4} \rightarrow \text{id}en$ 归约, 得

(1) (jnz, A, 0, 3) ;

(2) Expr $^{\wedge}$. FC \rightarrow (j, 0, 0, 0) ;

(3) Exprv. TC \rightarrow (jnz, A, 0, 0) ;

(4) (j, 0, 0, 5) ;

(5) Exprv. TC \rightarrow (jnz, C, 0, 0) ;

(6) (j, 0, 0, 7) ;

(7) Expr. TC \rightarrow (jnz, D, 0, 0) ;

(8) Expr. FC \rightarrow (j, 0, 0, 0) ;

- 当前句型 $\text{Expr}^{\wedge} (\text{Exprv} (\text{Exprv Expr}/\text{?F}))$, 用产生式 $\text{Expr}^{\wedge} \rightarrow \text{expr}' /\wedge'$ 归约, 得

(1) (jnz, A, 0, 3) ;

- (2) $\text{Expr}^{\wedge} \cdot \text{FC} \rightarrow (j, 0, 0, 0);$
- (3) $\text{Exprv. TC} \rightarrow (\text{jnz}, A, 0, 0);$
- (4) $(j, 0, 0, 5);$
- (5) $\text{Exprv. TC} \rightarrow (\text{jnz}, C, 0, 0);$
- (6) $(j, 0, 0, 7);$
- (7) $(\text{jnz}, D, 0, 9);$
- (8) $\text{Expr}^{\wedge} \cdot \text{FC} \rightarrow (j, 0, 0, 0);$

- 当前句型 $\text{Expr}^{\wedge} (\text{Exprv} (\text{Exprv} \text{ Expr}^{\wedge} ?F))$, 用产生式 $\text{Expr} \rightarrow \text{id}en$ 归约, 得

- (1) $(\text{jnz}, A, 0, 3);$
- (2) $\text{Expr}^{\wedge} \cdot \text{FC} \rightarrow (j, 0, 0, 0);$
- (3) $\text{Exprv. TC} \rightarrow (\text{jnz}, A, 0, 0);$
- (4) $(j, 0, 0, 5);$
- (5) $\text{Exprv. TC} \rightarrow (\text{jnz}, C, 0, 0);$
- (6) $(j, 0, 0, 7);$
- (7) $(\text{jnz}, D, 0, 9);$
- (8) $\text{Expr}^{\wedge} \cdot \text{FC} \rightarrow (j, 0, 0, 0);$

- (9) $\text{Expr. TC} \rightarrow (\text{jnz}, F, 0, 0);$
- (10) $\text{Expr. FC} \rightarrow (j, 0, 0, 0);$

- 当前句型 $\text{Expr}^{\wedge} (\text{Exprv} (\text{Exprv} \text{ Expr}^{\wedge} ?\text{Expr}))$, 用产生式 $\text{Expr} \rightarrow ?\text{Expr}$ 归约, 得

- (1) $(\text{jnz}, A, 0, 3);$
- (2) $\text{Expr}^{\wedge} \cdot \text{FC} \rightarrow (j, 0, 0, 0);$
- (3) $\text{Exprv. TC} \rightarrow (\text{jnz}, A, 0, 0);$

- (4) (j, 0, 0, 5) ;
- (5) Exprv. TC→ (jnz, C, 0, 0) ;
- (6) (j, 0, 0, 7) ;
- (7) (jnz, D, 0, 9) ;
- (8) Expr[^]. FC→ (j, 0, 0, 0) ;
- (9) Expr. FC→ (jnz, F, 0, 0) ;
- (10) Expr. TC→ (j, 0, 0, 0) ;

- 当前句型 Expr[^] (Exprv (Exprv Expr[^] Expr)), 用产生式 Expr→Expr[^] Expr 归约, 得

- (1) (jnz, A, 0, 3) ;
- (2) Expr[^]. FC→ (j, 0, 0, 0) ;
- (3) Exprv. TC→ (jnz, A, 0, 0) ;
- (4) (j, 0, 0, 5) ;
- (5) Exprv. TC→ (jnz, C, 0, 0) ;
- (6) (j, 0, 0, 7) ;
- (7) (jnz, D, 0, 9) ;
- (8) (j, 0, 0, 0) ;
- (9) Expr. FC→ (jnz, F, 0, 8) ;
- (10) Expr. TC→ (j, 0, 0, 0) ;

- 当前句型 Expr[^] (Exprv (Exprv Expr)), 用产生式 Expr→Exprv Expr 归约, 得

- (1) (jnz, A, 0, 3) ;
- (2) Expr[^]. FC→ (j, 0, 0, 0) ;
- (3) Exprv. TC→ (jnz, A, 0, 0) ;

- (4) (j, 0, 0, 5) ;
- (5) (jnz, C, 0, 0) ;
- (6) (j, 0, 0, 7) ;
- (7) (jnz, D, 0, 9) ;
- (8) (j, 0, 0, 0) ;
- (9) Expr. FC \rightarrow (jnz, F, 0, 8) ;
- (10) Expr. TC \rightarrow (j, 0, 0, 5) ;

- 当前句型 $\text{Expr}^{\wedge} (\text{Exprv}(\text{Expr}))$, 用产生式 $\text{Expr} \rightarrow (\text{Expr})$ 归约, 所得四元式序列不变
- 当前句型 $\text{Expr}^{\wedge} (\text{Exprv Expr})$, 用产生式 $\text{Expr} \rightarrow \text{Exprv Expr}$ 归约, 得

- (1) (jnz, A, 0, 3) ;
- (2) $\text{Expr}^{\wedge} . \text{FC} \rightarrow (j, 0, 0, 0) ;$
- (3) (jnz, A, 0, 0) ;
- (4) (j, 0, 0, 5) ;
- (5) (jnz, C, 0, 3) ;
- (6) (j, 0, 0, 7) ;
- (7) (jnz, D, 0, 9) ;
- (8) (j, 0, 0, 0) ;
- (9) Expr. FC \rightarrow (jnz, F, 0, 8) ;
- (10) Expr. TC \rightarrow (j, 0, 0, 5) ;

- 当前句型 $\text{Expr}^{\wedge} (\text{Expr})$, 用产生式 $\text{Expr} \rightarrow (\text{Expr})$ 归约, 所得四元式序列不变
- 当前句型 $\text{Expr}^{\wedge} \text{Expr}$, 用产生式 $\text{Expr} \rightarrow \text{Expr}^{\wedge} \text{Expr}$ 归约, 得

- (1) (jnz, A, 0, 3) ;
- (2) (j, 0, 0, 0) ;

- (3) (jnz, A, 0, 0) ;
 - (4) (j, 0, 0, 5) ;
 - (5) (jnz, C, 0, 3) ;
 - (6) (j, 0, 0, 7) ;
 - (7) (jnz, D, 0, 9) ;
 - (8) (j, 0, 0, 2) ;
 - (9) Expr. FC \rightarrow (jnz, F, 0, 8) ;
 - (10) Expr. TC \rightarrow (j, 0, 0, 5) ;
- 3) 假设所产生的四元式序列编号从 1 开始
1. 当前句型为 `while A<C&B>0 do if A=1 then C:=C+1 else while A<=D do A:=A+2` 用产生式 $W1 \rightarrow \text{while}$ 归约, 得
 - (1) $W1. \text{loop} \rightarrow$
 - 2. 当前句型为 $W1 \ A<C&B>0 \ \text{do if } A=1 \ \text{then } C:=C+1 \ \text{else while } A<=D \ \text{do } A:=A+2$ 用产生式 $\text{Expr} \rightarrow \text{id} \ \text{rop} \ \text{id}$ 归约, 得
 - (1) $W1. \text{loop} \rightarrow \text{Expr. TC} \rightarrow (j < A, C, 0) ;$
 - (2) $\text{Expr. FC} \rightarrow (j, 0, 0, 0) ;$
 - 3. 当前句型为 $W1 \ \text{Expr} \ \& B>0 \ \text{do if } A=1 \ \text{then } C:=C+1 \ \text{else while } A<=D \ \text{do } A:=A+2$ 用产生式 $\text{Expr}^{\wedge} \rightarrow \text{Expr}' \ \&'$ 归约, 得
 - (1) $W1. \text{loop} \rightarrow (j < A, C, 3) ;$
 - (2) $\text{Expr}^{\wedge}. \text{FC} \rightarrow (j, 0, 0, 0) ;$
 - 4. 当前句型为 $W1 \ \text{Expr}^{\wedge} \ B>0 \ \text{do if } A=1 \ \text{then } C:=C+1 \ \text{else while } A<=D \ \text{do } A:=A+2$ 用产生式 $\text{Expr} \rightarrow \text{id} \ \text{rop} \ \text{id}$ 归约, 得
 - (1) $W1. \text{loop} \rightarrow (j < A, C, 3) ;$
 - (2) $\text{Expr}^{\wedge}. \text{FC} \rightarrow (j, 0, 0, 0) ;$
 - (3) $\text{Expr. TC} \rightarrow (j >, B, 0, 0)$

(4) Expr2. FC \rightarrow (j, 0, 0, 0) ;

5. 当前句型为 $W1 \ Expr \ ^ \ Expr \ do \ if \ A=1 \ then \ C:=C+1 \ else \ while \ A \leq D \ do \ A:=A+2$ 用产生式 $Expr \rightarrow Expr \ ^ \ Expr$ 归约, 得

(1) W1. loop \rightarrow (j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) Expr. TC \rightarrow (j>, B, 0, 0)

(4) Expr. FC \rightarrow (j, 0, 0, 2) ;

6. 当前句型为 $W1 \ Expr \ do \ if \ A=1 \ then \ C:=C+1 \ else \ while \ A \leq D \ do \ A:=A+2$ 用产生式 $WED \rightarrow W1 \ Expr \ do$ 归约, 得

(1) WED. loop \rightarrow (j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j>, B, 0, 5)

(4) WED. CH \rightarrow (j, 0, 0, 2) ;

7. 当前句型为 $WED \ if \ A=1 \ then \ C:=C+1 \ else \ while \ A \leq D \ do \ A:=A+2$ 用产生式 $Expr \rightarrow iden \ rop \ iden$ 归约, 得

(1) WED. loop \rightarrow (j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j>, B, 0, 5)

(4) WED. CH \rightarrow (j, 0, 0, 2) ;

(5) Expr. TC \rightarrow (j=, A, 1, 0)

(6) Expr. FC \rightarrow (j, 0, 0, 0)

8. 当前句型为 $WED \ if \ Expr \ then \ C:=C+1 \ else \ while \ A \leq D \ do \ A:=A+2$ 用产生式 $Condition \rightarrow ifExpr \ then$ 归约, 得

(1) WED. loop \rightarrow (j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) ($j >$, B, 0, 5)

(4) WED. CH \rightarrow (j , 0, 0, 2) ;

(5) ($j =$, A, 1, 7)

(6) Condition. CH \rightarrow (j , 0, 0, 0)

9. 当前句型为 WED Condition C:=C+1 else while A<=D do A:=A+2 将赋值语句 C:=C+1 归约为 Statement, 得

(1) WED. loop \rightarrow ($j < A$, C, 3) ;

(2) (j , 0, 0, 0) ;

(3) ($j >$, B, 0, 5)

(4) WED. CH \rightarrow (j , 0, 0, 2) ;

(5) ($j =$, A, 1, 7)

(6) Condition. CH \rightarrow (j , 0, 0, 0)

(7) (+, C, 1, T1)

(8) (=, T1, 1, C)

10. 当前句型为 WED Condition Statement else while A<=D do A:=A+2 用产生式 CondStElse \rightarrow Condition Statement 归约,

(1) WED. loop \rightarrow ($j < A$, C, 3) ;

(2) (j , 0, 0, 0) ;

(3) ($j >$, B, 0, 5)

(4) WED. CH \rightarrow (j , 0, 0, 2) ;

(5) ($j =$, A, 1, 7)

(6) (j , 0, 0, 10)

(7) (+, C, 1, T1)

(8) (=, T1, 1, C)

(9) CondStElse \rightarrow (j, 0, 0, 0)

11. 当前句型为 WED CondStElse while A<=D do A:=A+2 用产生式 W1' \rightarrow while 归约, 得

(1) WED. loop \rightarrow (j < A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j >, B, 0, 5)

(4) WED. CH \rightarrow (j, 0, 0, 2) ;

(5) (j =, A, 1, 7)

(6) (j, 0, 0, 10)

(7) (+, C, 1, T1)

(8) (=, T1, 1, C)

(9) CondStElse \rightarrow (j, 0, 0, 0)

(10) W1. loop \rightarrow

12. 当前句型为 WED CondStElse W1 A<=D do A:=A+2 用产生式 Expr \rightarrow iden rop iden 归约, 得

(1) WED. loop \rightarrow (j < A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j >, B, 0, 5)

(4) WED. CH \rightarrow (j, 0, 0, 2) ;

(5) (j =, A, 1, 7)

(6) (j, 0, 0, 10)

(7) (+, C, 1, T1)

(8) (=, T1, 1, C)

(9) CondStElse \rightarrow (j, 0, 0, 0)

(10) W1. loop→Expr. TC→(j, A, D, 0)

(11) Expr. FC→(j, 0, 0, 0)

13. 当前句型为 WED CondStElse W1 Expr do A:=A+2 用产生式 WED→W1 Expr do 归约, 得

(1) WED. loop→(j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j>, B, 0, 5)

(4) WED. CH→(j, 0, 0, 2) ;

(5) (j=, A, 1, 7)

(6) (j, 0, 0, 10)

(7) (+., C, 1, T1)

(8) (=, T1, 1, C)

(9) CondStElse→(j, 0, 0, 0)

(10) WED. loop→(j, A, D, 12)

(11) WED. CH→(j, 0, 0, 0)

14. 当前句型为 WED CondStElse WED A:=A+2 将赋值语句归约 A:=A+2 为 Statement 得

(1) WED. loop→(j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j>, B, 0, 5)

(4) WED. CH→(j, 0, 0, 2) ;

(5) (j=, A, 1, 7)

(6) (j, 0, 0, 10)

(7) (+., C, 1, T1)

- (8) (=, T1, , C)
- (9) CondStElse → (j, 0, 0, 0)
- (10) WED. loop → (j, A, D, 12)
- (11) WED. CH → (j, 0, 0, 0)
- (12) (+, A, 2, T2)
- (13) (=, T2, , A)

15. 当前句型为 **WED CondStElse WED Statement**, 用产生式 **Statement → WED Statement** 归约, 得

- (1) WED. loop → (j < A, C, 3) ;
- (2) (j, 0, 0, 0) ;
- (3) (j >, B, 0, 5)
- (4) WED. CH → (j, 0, 0, 2) ;
- (5) (j =, A, 1, 7)
- (6) (j, 0, 0, 10)
- (7) (+, , C, 1, T1)
- (8) (=, T1, , C)
- (9) CondStElse → (j, 0, 0, 0)
- (10) (j, A, D, 12)
- (11) Statement. CH → (j, 0, 0, 0)
- (12) (+, A, 2, T2)
- (13) (=, T2, , A)
- (14) (j, 0, 0, 10)

16. 当前句型为 **WED CondStElse Statement**, 用产生式 **Statement → CondStElse Statement** 归约, 得

(1) WED. loop \rightarrow (j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j>, B, 0, 5)

(4) WED. CH \rightarrow (j, 0, 0, 2) ;

(5) (j=, A, 1, 7)

(6) (j, 0, 0, 10)

(7) (+, C, 1, T1)

(8) (=, T1, , C)

(9) (j, 0, 0, 0)

(10) (j, A, D, 12)

(11) Statement. CH \rightarrow (j, 0, 0, 9)

(12) (+, A, 2, T2)

(13) (=, T2, , A)

(14) (j, 0, 0, 10)

17. 当前句型为 WED Statement, 用产生式 Statement \rightarrow WED Statement 归约, 得

(1) (j<A, C, 3) ;

(2) (j, 0, 0, 0) ;

(3) (j>, B, 0, 5)

(4) Statement. CH \rightarrow (j, 0, 0, 2) ;

(5) (j=, A, 1, 7)

(6) (j, 0, 0, 10)

(7) (+, C, 1, T1)

(8) (=, T1, , C)

(9) (j, 0, 0, 1)

(10) (j, A, D, 12)

(11) . (j, 0, 0, 1)

(12) (+, A, 2, T2)

(13) (=, T2, , A)

(14) (j, 0, 0, 10)

(15) (j, 0, 0, 1)

(16)

5.9 解：

1) First \rightarrow for iden ':=' Expr

```
{ int T=NewTemp();
  GEN(=, $4. PLACE, 0, ENTRY($2));
  $$. VAR= ENTRY($2);
  GEN(=, 0, , T); $$. temp=ENTRY(T);
  $$. loop=NXQ; }
```

Second \rightarrow First step Expr

```
{ int U=NewTemp();
  $$. loop=$1. loop;
  GEN(j=, $1. temp, 0, NXQ+2);
  GEN(+, $1. VAR, $3. PLACE, $1. VAR);
  GEN(=, 1, , $1. temp);
  $$. VAR=$1. VAR; }
```

\$\$. DELTA=\$3. PLACE; }

Third→Second until Expr

{ int T, T1, q=NQ;

\$\$. loop=\$1. loop;

T=NewTemp();

GEN(J<, \$1. DELTA, 0, q+4);

GEN(J>, \$1. DELTA, 0, q+6);

GEN(=, 0, , T);

GEN(J, , , q+7);

GEN(=, -1, , T);

GEN(J, , , q+7);

GEN(=, 1, , T);

T1=NewTemp();

GEN(-, \$1. DELTA, \$3. PLACE, T1);

GEN(*, T, T1, T1);

GEN(J<, T1, 0, q+11);

\$\$. CH=NQ;

GEN(J, , , 0);

S→ Third do S

{ BackPatch(\$3. CH, \$1. loop);

GEN(J, , , \$3. loop);

\$\$. CH=\$3. CH; }

2)

- $R \rightarrow \text{repeat}\{\$\$.loop = \text{NXQ};\}$
- $Rs \rightarrow R S \quad \{\$\$.loop = \$1. \text{loop}; \text{BackPath}(\$2. \text{chain}, \text{NXQ});\}$
- $S \rightarrow Rs \text{ until Expr } \{\$\$.chain = \$3. \text{TC}; \text{BackPatch}(\$3. \text{FC}, \$1. \text{loop});\}$

5.10 解:

1) (*, I, 20, T1)

- (+, J, T1, T1)
- (-, a, Ca, T2)
- (=, J, 0, T)
- (+, T, 1, T)
- (=, I, 0, P1)
- (=, J, 0, P2)
- (-, P1, P2, P1')
- (=, I, 0, T1)
- (=, J, 0, T2)
- (+, T1, T2, T1')
- (*, P1', 20, T3')
- (+, T1', T3', T3')
- (-, a, Ca, T')
- (=[], T' [T3'], T3)
- (*, T, 20, T)
- (+, T3, T, T)
- (=, I, 0, T4)
- (+, T4, 1, T4)
- (*, T, 10, T5)
- (+, T4, T5, T5)
- (-, b, Cb, T6)
- (=[], T6[T5], 0, T7)
- ([]=, T7, 0, T2[T1])

5.11 解:

- 采用值调用:N=2
- 采用引用调用:N=9

5.12 解:

- 采用值调用:b[]={1, 2, 3, 4}
- 采用引用调用:b[]={20, 5, 3, 4}
- 采用结果调用:error
- 采用值结果调用:b[]={3, 5, 3, 4}

5.13 解：

- 1) DPart → var D
- 2) D → VarList ' : ' Type { BackChain(\$1. CH, \$2. type); }
- 3) Type → integer { \$\$. type = ' i' ; }
- 4) Type → real { \$\$. type = ' r' ; }
- 5) Type → boolean { \$\$. type = ' b' ; }
- 6) Type → char { \$\$. type = ' c' ; }
- 7) VarList → iden { \$\$. CH = NewChain(); AddToChain(\$\$. CH, ENTRY(\$1)); }
- 8) VarList → VarList ',' iden
{ AddToChain(\$1. CH, ENTRY(\$1)); \$\$. CH = \$1. CH; }

5.14 解：源程序

```
PROGRAM ProcAsPara;  
USES wincrt;  
VAR  
    putin, result:integer;  
PROCEDURE child(VAR a: INTEGER, FUNCTION f, i:INTEGER);  
BEGIN a:=f(i) END;  
FUNCTION func( x: INTEGER) : INTEGER;  
BEGIN func:= x*x; RETURN END;  
BEGIN  
    putin:=8;  
    child(result, func, putin);  
    writeln(result);
```

END.

执行结果: 64

5.15 解:

- 由于 C 语言数值的下界为 0, 所以内情向量中只需放各维的上界, 即放 **ui, n, c, a**, 共 $n+3$ 个单元;
-
- **Varable** → **ArrayVar** {**\$\$=\$\$1}**
- **ArrayMSG** → **iden** [**number**

{

\$\$=Entry(\$1);

VarList[\$\$]. CAT=Array; /*种属为数组*/

VarLIst[\$\$]. IsPointer=0;

VarList[\$\$]. ADDR->DIM=1; /*记录维数*/

/*下面为内情向量申请空间, 并填入第一维下标信息, 其中, 前两个单元(下标为[0]和[1])用来存放 a、c 之值(此时暂不填写), n 值可由 DIM 保存, 因此不必另存。*/

VarList[\$\$]. ADDR->Vector=malloc(3*sizeof(int));

VarList[\$\$]. ADDR->Vector[2]=\$3; /*第一维上界*/

}

|ArrayMSG , number

{int dim= VarList[\$\$]. ADDR->DIM+1;

\$\$=\$\$1;

VarList[\$\$]. ADDR->DIM++; /*维数加 1*/

/*下面增加向量空间, 记录新一维的信息*/

VarList[\$\$]. ADDR->Vector=

realloc(VarList[\$\$]. ADDR->Vector,

```
(dim+2)*sizeof(int));  
/*下面记录当前维的上界*/  
VarList[$$].ADDR->Vector[3*dim-1]=$3;  
}  
ArrayVar → ArrayMSG ]  
{  
$$=$1; /*传递数组名在表中序号*/  
FillArrMSG_C($$);  
/*计算并填写数组内情向量的 C 值*/  
}
```

5.16 解：在此种情况下，可以通过使用堆栈，从左到右依次处理各下标表达式，且每当处理完一个下标表达式 E_i 时，就将相应的 $E_i.PLACE$ 推入堆栈，待全部下标表达式处理完毕之后，再产生按从右到左累计 VARPART 的四元式序列。这需要在有关的语义子程序中用一段循环程序来实现。属性翻译文法略。

5.17 解：

- Condition → if (Expr)

```
{BackPatch($2. TC, NXQ); $$.Chain=$2. FC;}
```

- Statement → Condition Statement

```
{$$.Chain=Merge($1. Chain, $2. Chain);}
```

- CondStElse → Condition Statement ; else

```
{int q=NXQ;
```

```
GEN(j, 0, 0, 0);
```

```
BackPatch($1. Chain, NXQ);
```

```
$$.Chain=Merge($2. Chain, q);}
```

- Statement→CondStElse Statement ;

```
{$$.Chain=Merge($1.Chain, $2.Chain);}
```

第六章 习题解答

1. 答：对于有嵌套分程序结构的程序设计语言的编译程序而言，可将 \sin 、 \cos 等标准函数的名字视为最外层分程序定义的全局变量名，这样对其的任何引用均可在最外层符号表中找到其相关信息。另一种解决方案是，在文法定义中就将其定义为一类特殊（类似于关键字）的终结符，并为其定义专门的调用标准函数的文法，这样在词法分析阶段就可区分出它与其它标识符的不同。

2. 解: (1) 当扫描到 PROCEDURE a 时, 符号表如图 6-1:

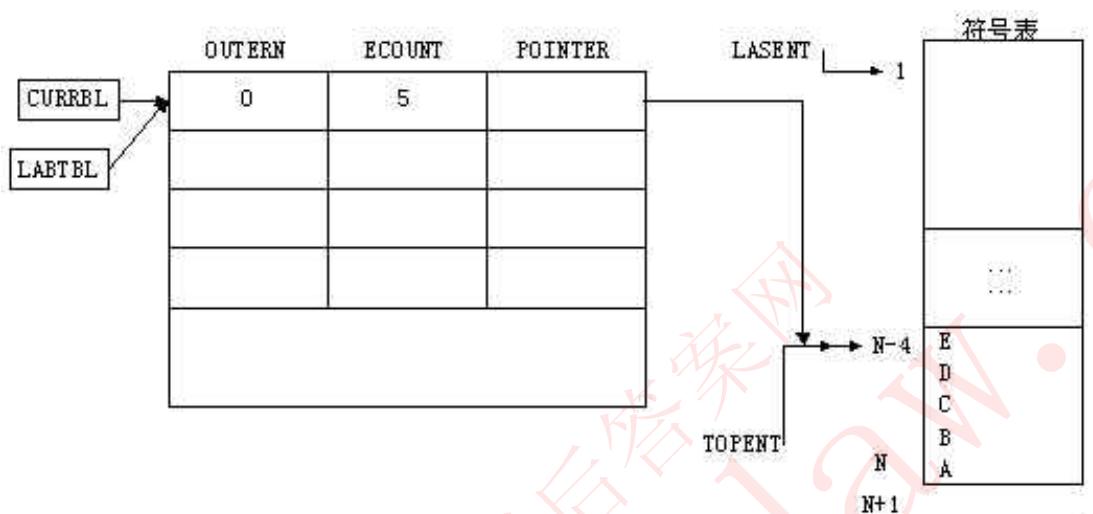


图 6-1-7

(2) 当扫描到 PROCEDURE b 所在行时, 符号表如图 6-2。

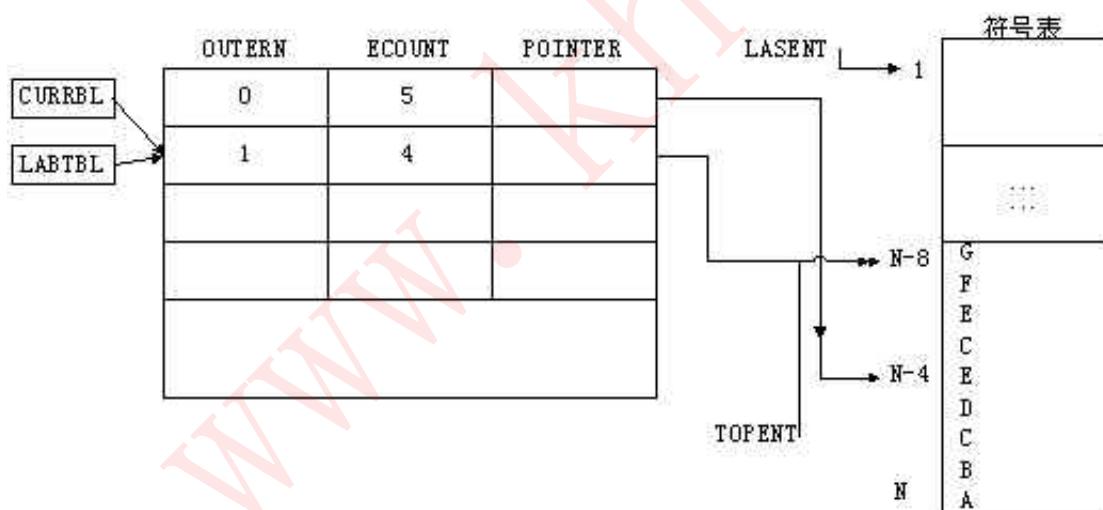


图 6-2

(3) 当扫描到 PROCEDURE c 所在行时, 符号表如图 6-3。

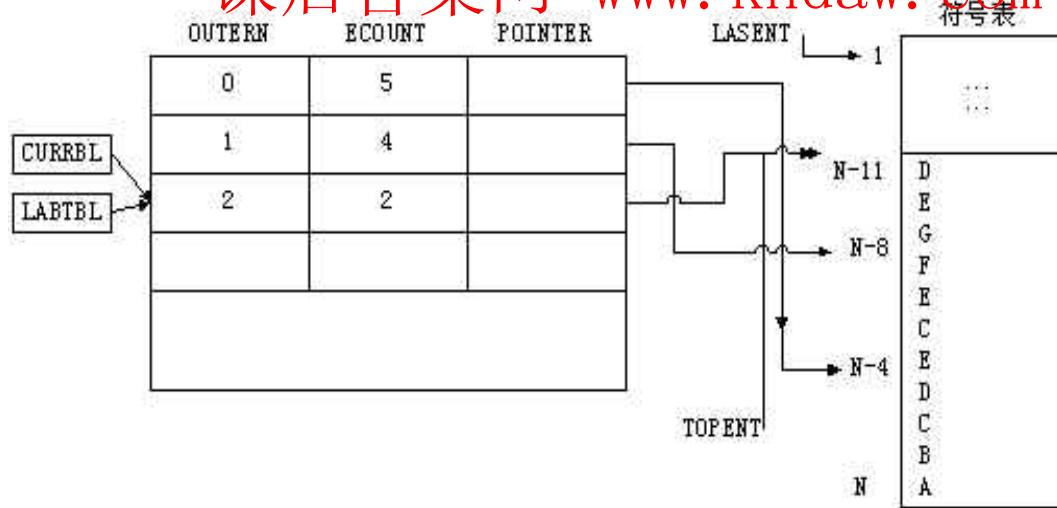


图 6-3

(4) 当扫描到 PROCEDURE c 的 begin 语句时, 符号表如图 6-4。

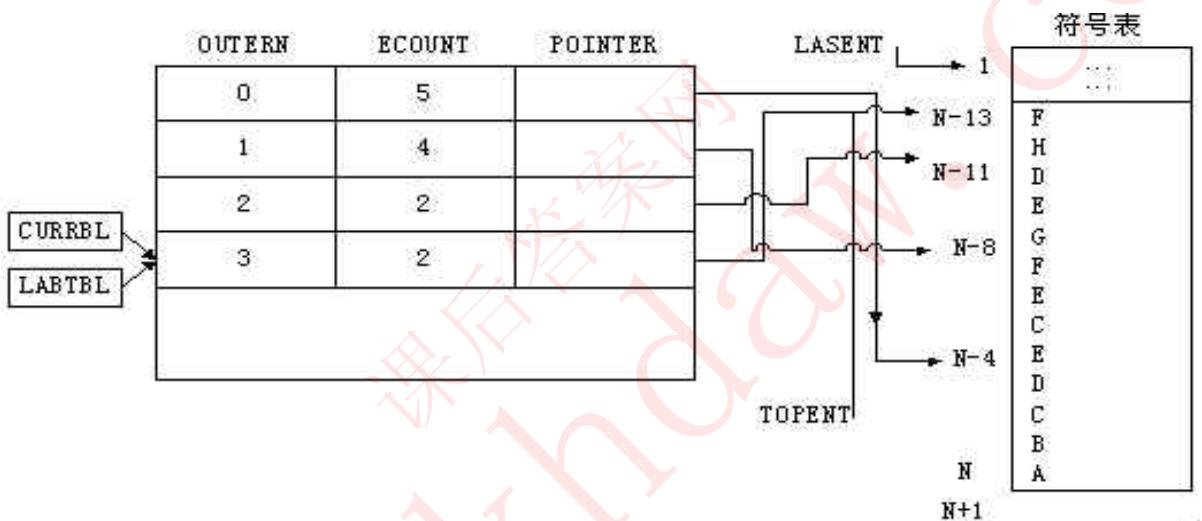


图 6-4

(5) 当扫描到 PROCEDURE b 的 begin 语句时, 符号表如图 6-5。

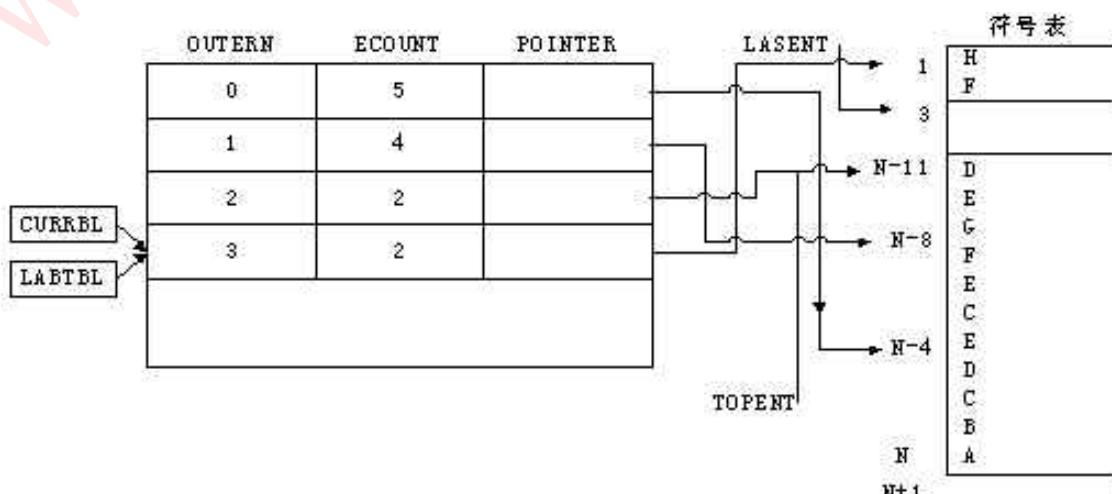


图 6-5

(6) 当扫描到 PROCEDURE A 的 begin 语句时, 符号表如图 6-6。

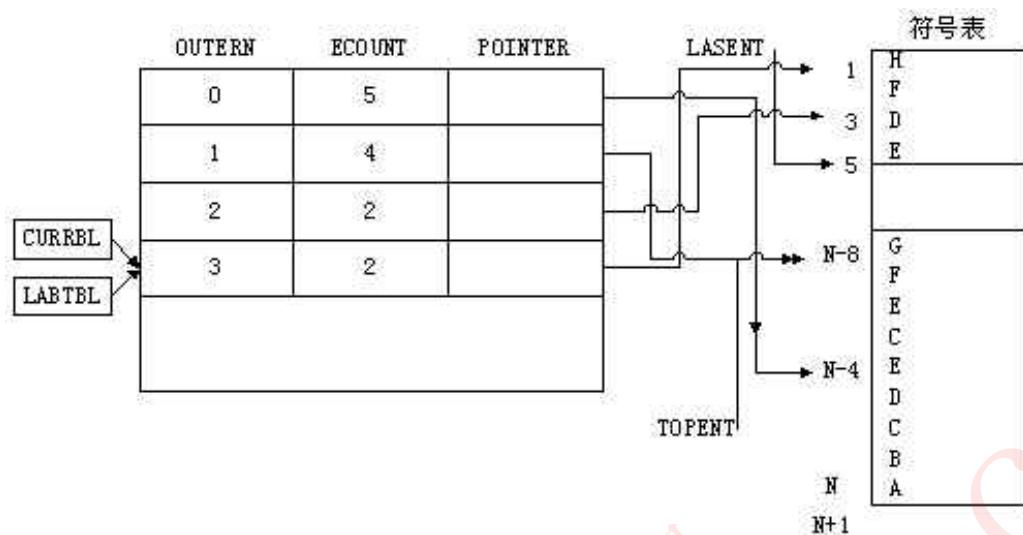


图 6-6

(7) 当扫描到 PROGRAM ex62 的 begin 语句时, 符号表如图 6-7。

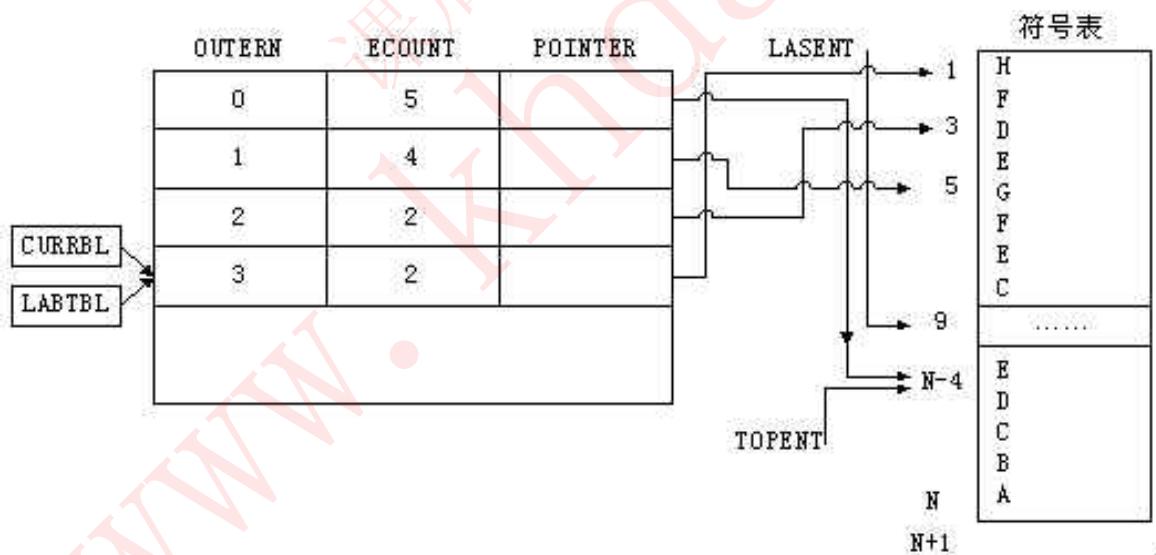


图 6-7

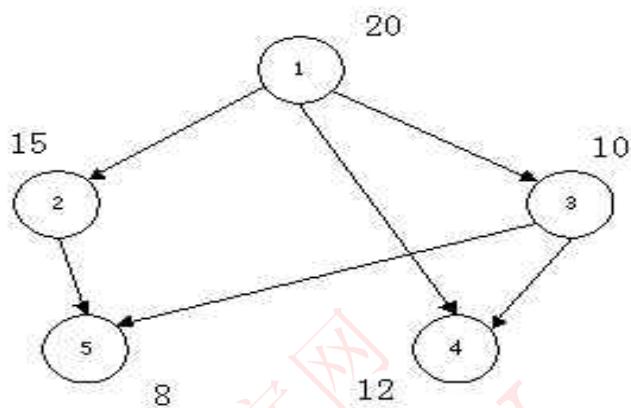
(8) 因主程序外层再无其它程序, 所以, 主程序的变量在符号表中的位置不必再挪动。也就是说, 图 6-7 即为最终符号表的状态。

3. 略

www. Khdaw. Com

第七章 习题解答

7. 1 设有由五个程序段组成的 FORTRAN 程序，各程序段间的相互调用关系如下图所示。试为此程序的各程序段合理地分配局部数据区。



解：假定数据区从 1 号单元开始分配，则各程序段所占单元情况为：

5: $1^{\sim}8$ 4: $1^{\sim}12$

2: $9^{\sim}23$ 3: $13^{\sim}22$

1: $24^{\sim}44$

整个程序占用单元数为 44

7. 2 解：（提示）一般来说，FORTRAN 语言不允许递归调用，但嵌套调用是允许的。其产生中间代码的方式与其它语言别无二致。

7. 3 解：设每个变量均占用一个单元，临时变量的个数 $\leq m$ 。定义一个数组 $a[m]$ 记录每个临时变量（按出现顺序编号命名）分配单元的地址 ($\leq n \leq m$, 按 1、2、3... 编号)，初值均为 0。用另一数组 $T[n]$ 标记每个单元被使用的最后期限，初值为 0。引入全局变量 $LAST$ 记录当前被使用的最后一个单元的地址，初值为 1。

- 令 $i=1$ ；
- 取出有序对 (F_i, L_i) ；
- 令 $j=1$ ；
- 若 $(T[j] < F_i) \wedge (T[j] \neq 0)$ ，

则令: $T[j] = L_i; a[i] = j$; 即将第 j 单元分配给变量 i , $T[j]$ 中存放 i 的失效时间 L_i 。转 7) ;

否则, $j++$;

- $j \leqslant LAST$, 转 4) ;
- 为变量 i 分配单元: 令 $T[LAST] = L_i; a[i] = LAST$; $LAST++$;
- $i++$; 若 $i \leqslant m$, 转 2) ; 否则, 终止。

7. 4 解: 假定在运算过程中每个临时变量的值只使用一次, 即当其被用作运算对象后便成为无用量。则可按下述方法计算临时变量的使用个数:

- 引入变量 `int TmpVarNum` 记录所求临时单元的最小个数, 当表达式计算到当前位置时占用的临时单元数用 `int CurTmpVarNum` 记录, 初始值都是 0;
- 在处理表达式过程中, 目前正在运算的两个对象有以下三种情况:

①若两个运算对象均是用户定义的变量或常数, 则必须引入临时变量存放运算结果:

```
CurTmpVarNum++;
if (CurTmpVarNum > TmpVarNum) TmpVarNum = CurTmpVarNum;
```

②若两个运算对象均为临时变量, 则只需用其中之一存放运算结果, 另一临时变量可释放之: `CurTmpVarNum--`;

③若运算对象之一为临时变量, 则只需用该临时变量存放运算结果即可, 使用临时变量的个数不变。

5. 解: 为便于描述, 我们用过程名代表其相应的活动记录, 并视主程序 `ex75` 亦为一过程 (0 层)。则程序运行至各标号处的运行栈情况如下:

1) 运行至 L6 处:

`ex75`

2) 运行至 L4 处:

`ex75` `p2`

- 运行至 L5 处与至 L4 处相同;
- 运行至 L1 处:

`ex75` `P2` `P1`

- 运行至 L3 处与至 L5 处相同;
- 运行至 L2 处:

ex75 P2 P1 F

- 从函数 F 返回时:

ex75 P2 P1

8) 从 P1 返回时:

ex75 P2

9) 从 P2 返回时直至 L7 处时:

ex75

6. 解: 与上题描述方法相同:

- 调用 f(10) 后, 栈内容为:

ex76 f(10)

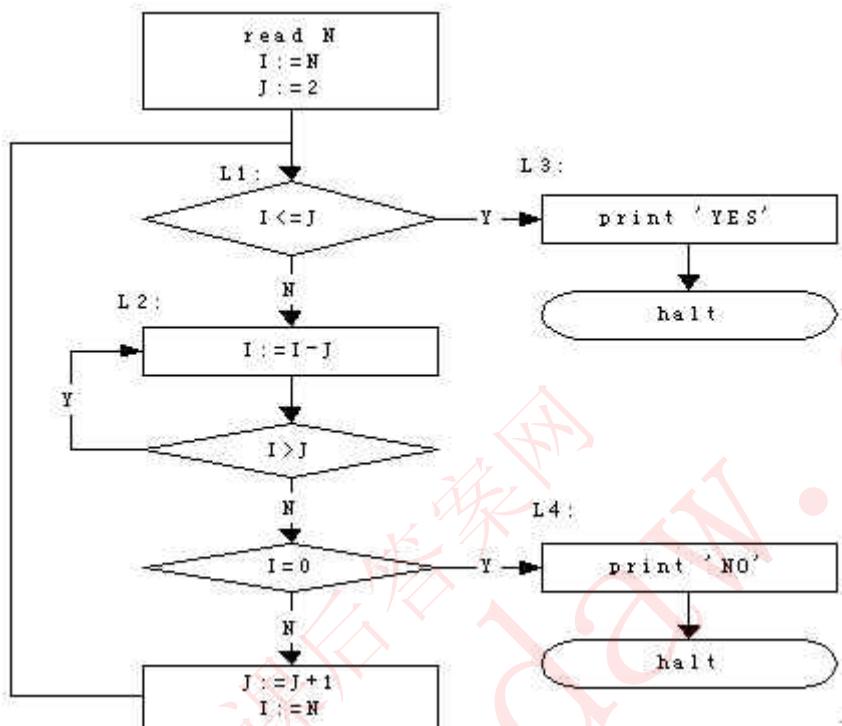
2) 在函数 f() 内, 又一次调用了函数 f 自身 (即第二次进入 f) :

ex76 f(10) f(9)

7. 略

第八章 习题解答

8.1 解：划分情况及控制流程如下图：



8.2 解：四元式序列如下（省略临时变量使用及形实结合代码）：

```

goto prog

lowterm: numcopy:=num

dencopy:=den

loop:  if dencopt<>0 goto L1

      goto L2

L1:   remainder:=numcopy mod dencopy

      numcopy:=dencopy

      dencopy:=remainder

      goto loop
  
```

L2: if numcopy>1 goto L3

goto L4

L3: num:=num div numcopy

den:=den div numcopy

L4: return

prog: input numerator

input denominator

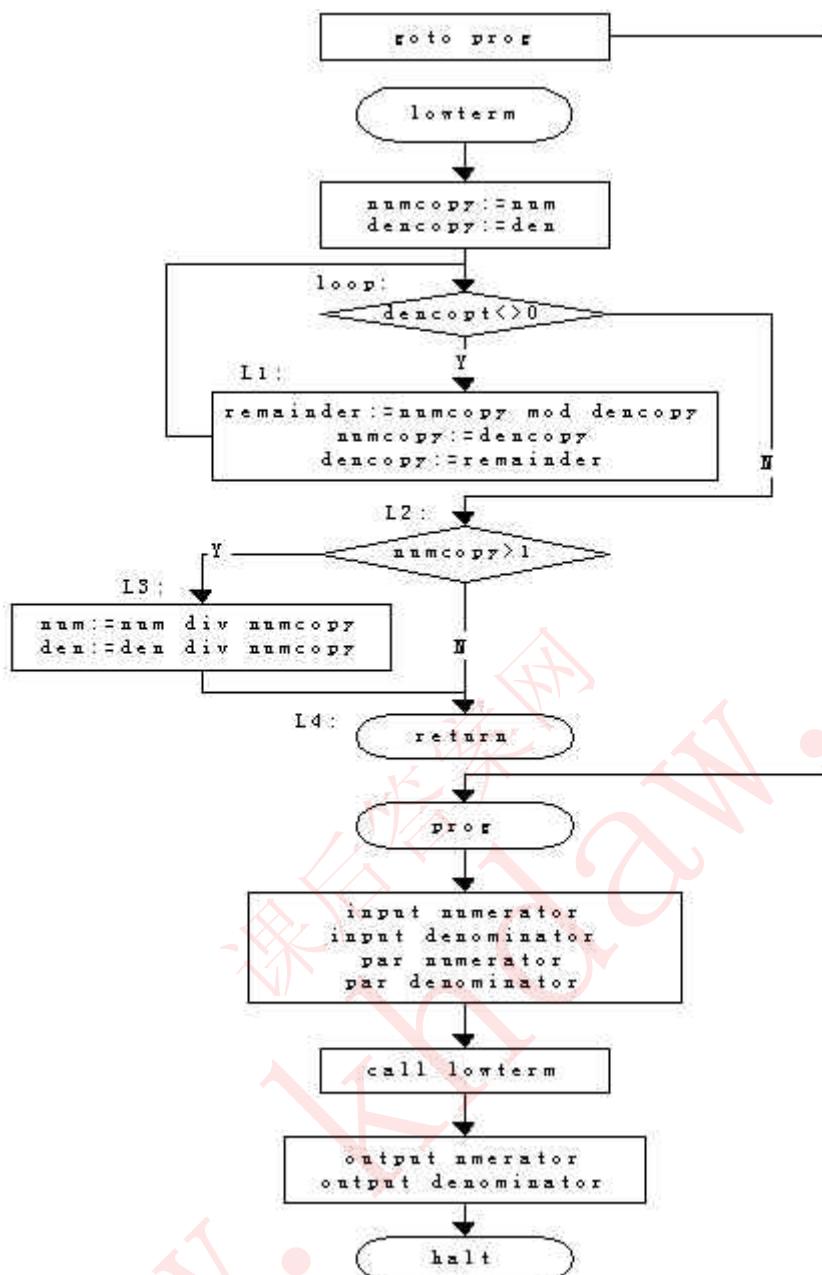
par numerator

par denominator

call lowterm

output nmerator

output denominator



8.3 解：四元式序列为：

```

i:=1
goto CHECKi

LOOPi: i:=i+1

CHECKi: if i<=n goto L1

        goto OUTi
    
```

```
L1:      j:=1
        goto CHECKj
LOOPj:   j:=j+1
CHECKj:  if j<=n goto L2
        goto OUTj
L2:      T1:=i-1
        t1:=T1*n
        T1:=T1+j
        T2:=addr(C)
        T1[T2]:=0
        goto LOOPj
OUTj:    goto LOOPi
OUTi:    i:=1
        goto CHKi
LPi:     i:=i+1
CHKi:   if i<=n goto L3
        goto OTi
L3:     j:=1
        goto CHKj
LPj:     j:=j+1
CHKj:   if j<=n goto L4
        goto OTj
L4:     k:=1
```

goto CHKk

LPk: k:=k+1

LLk: if k<=n goto L5

goto OTk

L5: T1:=i-1

T1:=T1*n

T1:=T1+j

T2:=addr (C)

T3:=i-1

T3:=T3*n

T3:=T3+j

T4:=addr (C)

TT1:=T3[T4]

T5:=i-1

T5:=T5*n

T5:=T5+k

T6:=addr (A)

TT2:=T5[T6]

T7:=k-1

T7:=T7*n

T7:=T7+j

T8:=addr (B)

TT3:=T7[T8]

T9:=TT2*TT3

T10:=TT1+TT2

T1[T2]:=T10

goto LPk

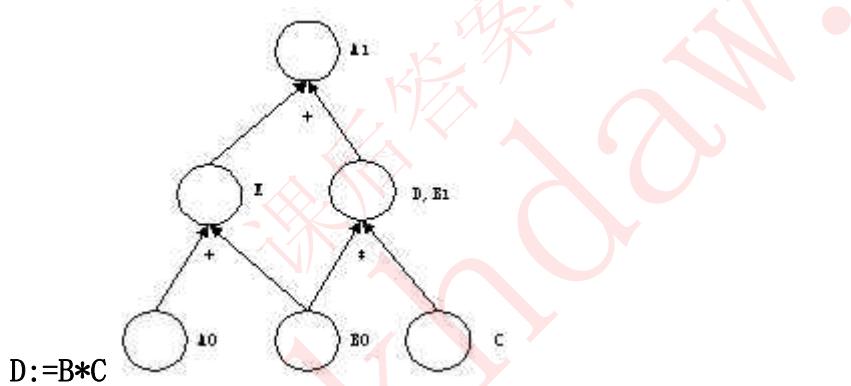
0Tk: goto LPj

0Tj: goto LPi

0Ti: halt

基本块划分与流程图略

8.4 解: DAG图见右, 优化后的代码为



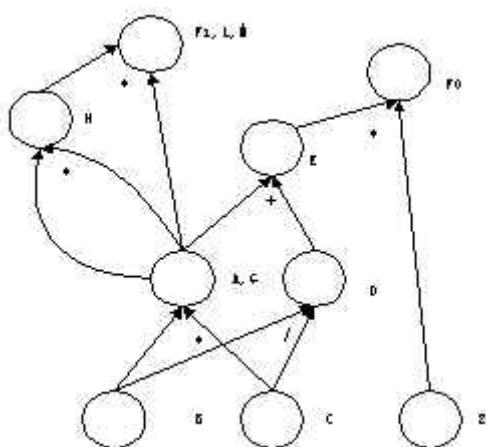
E:=A+B

B:=D

A:=E+D

8.5 解:

(1) DAG 图见下图。



若只有 G、L、M 在出口之后活跃，则优化

后的代码为：

$G := B * C$

$H := G * G$

$L := H * G$

$M := L$

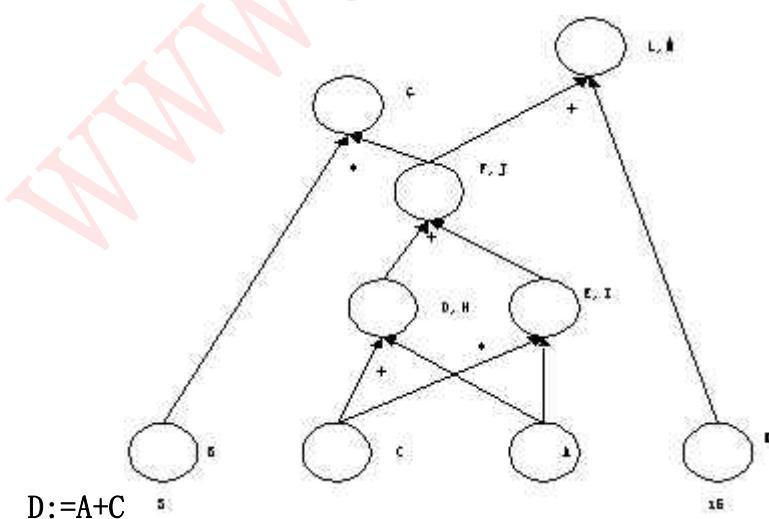
若只有 L 在出口之后活跃，则代码为

$G := B * C$

$H := G * G$

$L := H * G$

(2) DAG 见右图。若只有 G、L、M 活跃，则代码为



E:=A*C

F:=D+E

G:=3+F

L:=F+15

M:=L

若只有 L 活跃, 则代码为

D:=A+C

E:=A*C

F:=D+E

L:=F+15

8. 6 解: (a) 必经结点集: D2={2} D3={2, 3}, D4={2, 4} D5={2, 4, 5}
D6={2, 4, 6} D7={2, 4, 7} D8={2, 4, 7, 8}

回边及相应循环: 7→4: {4, 5, 6, 7}

8→2: {2, 3, 4, 5, 6, 7, 8}

(b) 必经结点集: D1={1} D2={1, 2} D3={1, 2, 3} D4={1, 2, 3, 4} D5={1, 2, 3, 5}
D6={1, 2, 3, 6} D7={1, 2, 7} D8={1, 2, 7, 8}

回边及相应循环: 7→2: {2, 3, 4, 5, 6, 7}

(c) 必经结点集: D1={1}, D2={1, 2}, D3={1, 2, 3}, D4={1, 2, 4}, D5={1, 2, 5},
D6={1, 2, 3, 6}, D7={1, 2, 7}

回边及相应循环: 5→2: {2, 3, 4, 5}

6→6: {6}

注意: 5→4 不是回边。因为 4 不是 5 的控制结点。

8. 7 略

8. 8 解

(1) 四元式序列:

```
I:=1
      goto CHECK
LOOP:   I:=I+1
CHECK:  if I<=n goto L
      goto OUT
L:      T1:=I-1
      T1:=T1*n
      T1:=T1+J
      T2:=addr(A)
      T3:=I-1
      T3:=T3*n
      T3:=T3+J
      T4:=addr(A)
      T5:=T3[T4]
      T6:=J-1
      T6:=T6*n
      T6:=T6+I
      T7:=addr(A)
      T8:=T6[T7]
      T9:=T7+T8
      T1[T2]:=T9
      goto LOOP
```

OUT: halt

(2) 无循环不变量;

(3) 优化后代码:

I:=1

goto CHECK

LOOP: I:=I+1

CHECK: if I<=n goto L

goto OUT

L: T1:=I-1

T1:=T1*n

T1:=T1+J

T2:=addr(A)

T3:=T1[T2]

T6:=J-1

T6:=T6*n

T6:=T6+I

T6:=T6[T2]

T3:=T3+T8

T1[T2]:=T3

goto LOOP

OUT: halt

8.9 优化后的代码为

read J, K

A:=0

B:=0

I:=100*K

L: A:=A+K

B:=B+J

C:=A*B

write C

if A<I goto L

halt

10. \sim 13 略