

编译原理 14 年

西安电子科技大学

考试时间 120 分钟

试 题 A

题号	一	二	三	四	五	六	七	八	九	十	总分
分数											

1. 考试形式：闭卷；2. 考试日期：2014 年 月 日 3. 本试卷共 大题，满分 100 分。

班级 _____ 学号 _____ 姓名 _____ 任课教师 _____

注意：所有题目请写在答题纸上。

一、填空题（每空 1 分，共 15 分）

1. 词法分析 和 语义分析 是编译程序各阶段都涉及到的工作。
2. 在编译器工作过程中，实现语言关键字大小写不敏感的阶段是 词法分析，分析语言结构的阶段是 语法分析。
3. 不确定有限自动机中的有限是指 状态 的数量是有限的。
4. 程序的语义错误可分为 静态 语义错误和 动态 语义错误。
5. 推导的过程可以用一棵树来表示，被称为 解析树。
6. $a * -(b + c)$ 的后缀式是 $abc+-*$ 。
7. 程序运行时内存的划分与数据空间的动态存储分配策略有 栈空间、堆空间。
8. 当文法的 左递归及右递归 两两 不相交 时，该文法对应的句子的分析不含有 回溯性。
9. 在自上而下的语法分析方法中，应先消除文法的 左递归 递归，再消除文法的 右递归 递归。

二、简答题（每题 5 分，共 25 分）

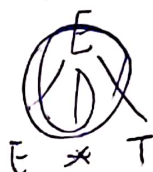
1. 解释什么是 LR(0) 项目集中的移进/归约冲突，简单说明 SLR(1) 方法如何解决移进/归约冲突。

地址：新综一楼辉扬快印

2. 为什么在一般情况下用正规式而不用 CFG 来描述语言的词法?
3. 符号表的作用是什么? 符号表上的操作主要有哪些?
4. 控制栈中的活动记录保存的信息有什么? 具体内容有什么?
5. 按照文法 $G: E \rightarrow E * T \mid T$

$$T \rightarrow T + a \mid a$$

句子 $2 + 3 * 4 + 5$ 按照文法规约, 其值为多少;
该文法为二义文法吗, 为什么?



三、计算题 (共 60 分)

1. (12 分) 已知正规式 $01(0|1)^*10$ 描述的正规集, 试给出:

- (1) (4 分) 识别该正规集的 NFA;
- (2) (4 分) 识别该正规集的 DFA (要有计算过程);
- (3) (4 分) 最小化的 DFA。

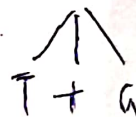


2. (9 分) 设文法 $G: E \rightarrow E + T \mid T$

$$T \rightarrow T * P \mid P$$

$$P \rightarrow (E) \mid j$$

给出句型 $P + (T + j)$ 的短语、直接短语及句柄。



3. (12 分) 已知文法 G 为: $E \rightarrow aTFb$ $T \rightarrow b \mid Tba$ $F \rightarrow c$

- (1) (4 分) 改写 G 为等价的 LL(1) 文法;
- (2) (4 分) 求每个非终结符的 FIRST 集合和 FOLLOW 集合;
- (3) (4 分) 构造预测分析表。

4. (12 分) 已知文法 G 为: $E \rightarrow E * T \mid T$ $T \rightarrow F + T \mid F$ $F \rightarrow id$

- (1) (4 分) 画出句子 $a * b + c + d$ 的分析树;
- (2) (4 分) 给出其识别活前缀的 DFA;
- (3) (4 分) 该文法是 SLR(1) 文法吗? 为什么?

5. (7 分) 将语句 if $a < c$ or $b > d$ then while $a > 3$ and $b < 4$ do $a := a + b$;
翻译成三地址码, 其中 a 、 c 为整型, b 、 d 为实型, 布尔表达式要求采用短路计算。

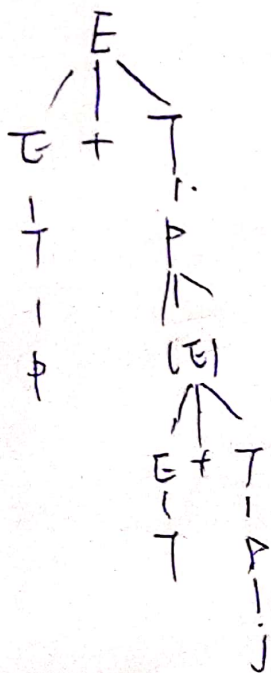
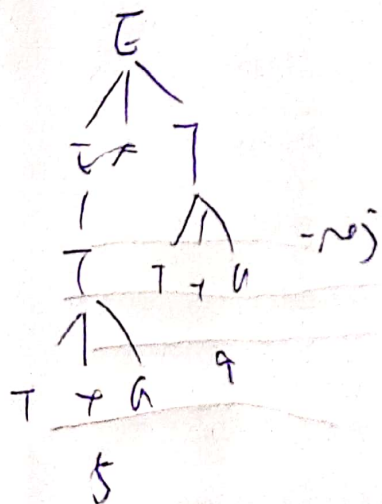
6. (8 分) 有一过程 A 如下所示。采用静态作用域、最近嵌套原则, 设 A 是第 0 层的过程。

procedure A is

 procedure B is $x : \text{integer}; \text{begin} \dots \text{end } B;$


```
begin ..... end A;
```

(2) (4 分) 若程序运行控制流是 A—D—E—B—C, 给出控制栈中的内容和控制链与访问链的指向。


$$Z_{12} : p, T, j, T^j, (T^j) \cdot p + 1(T^j).$$

Sign: P.T.J.

12/12 : p.

一、 填空题 (每题 1 分, 共 15 分)

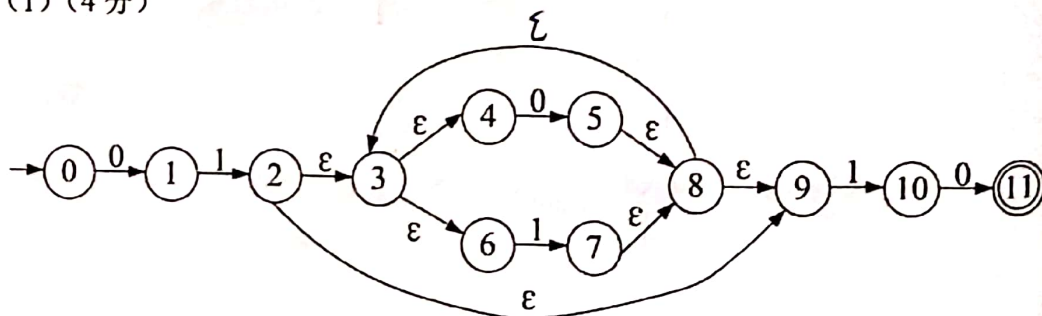
1. 符号表管理、出错处理
2. 词法分析、语法分析
3. 状态
4. 静态、动态
5. 分析树
6. $abc+^*$
7. 栈分配、堆分配
8. 候选项首符集、不相交
9. 直接左、间接左

二、 简答题 (每题 5 分, 共 25 分)

1. 移进/归约冲突是指在一个 LR(0) 项目集中同时存在可移进项 $A \rightarrow \beta 1, \beta 2$ 和可归约项 $B \rightarrow \beta 1$ 。对于移进/归约冲突, 计算非终结符 B 的 FOLLOW(B) 集合并看是否 $FOLLOW(B) \cap FIRST(\beta 2) = \Phi$, 若为 Φ 则可解决冲突, 否则不可解决。
2. 语法规则简单, 用正规式描述已足够; 正规式的表示比 CFG 更直观、简洁、易于理解; 有限自动机的构造比下推自动机简单, 且分析效率高; 区分词法和语法, 为编译器前端的模块划分提供方便。
3. 符号表的作用: 连接声明与引用的桥梁, 记住每个符号的相关信息, 如作用域和绑定等, 帮助编译的各个阶段正确有效地工作。符号表上的操作主要有: 查找、插入、删除、修改。
4. 活动记录保存的有控制信息和访问信息, 具体内容有: 参数与返回值、控制链(可选)、访问链(可选)、调用时需要保存的机器状态、过程内部声明的数据、临时变量。
5. 45. 不是二义文法, + 优先级比 * 高, + 和 * 都是左结合。

三、 计算题

1. (12 分)
(1) (4 分)



(2) (4 分)

ε -闭包($\{0\}$)= $\{0\}$ A

ε -闭包(smove(A, 0))= $\{1\}$ B

ε -闭包(smove(B, 1))= $\{2,3,4,6,9\}$ C

ε -闭包(smove(C, 0))= $\{5,8,9,3,4,6\}$ D

ε -闭包(smove(C, 1))= $\{7,8,9,10,3,4,6\}$ E

ε -闭包(smove(D, 0))= $\{5,8,9,3,4,6\}$ D

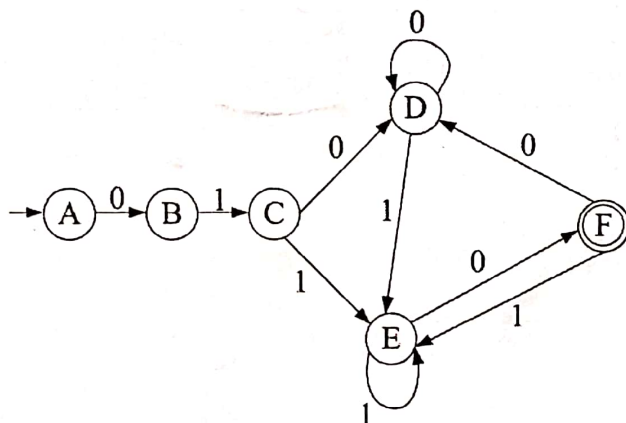
ε -闭包(smove(D, 1))= $\{7,8,9,10,3,4,6\}$ E

ε -闭包(smove(E, 0))= $\{11,5,8,9,3,4,6\}$ F

ε -闭包(smove(E, 1))= $\{7,8,9,10,3,4,6\}$ E

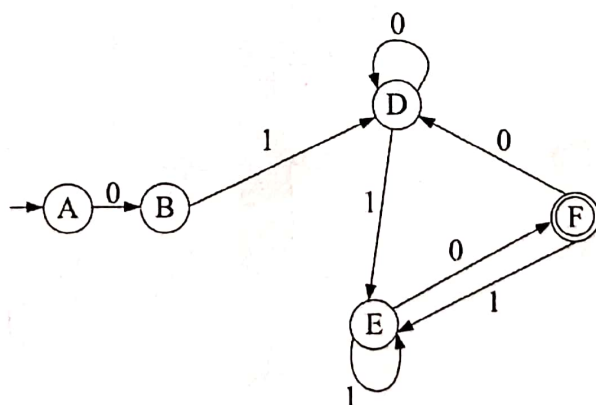
ε -闭包(smove(F, 0))= $\{5,8,9,3,4,6\}$ D

ε -闭包(smove(F, 1))= $\{7,8,9,10,3,4,6\}$ E



(3) (4 分)

初始化划分 $\Pi_1 = \{ABCDE, F\}$, $\Pi_{final} = \{A, B, CD, E, F\}$



2. (9 分)

短语: P, T, j, T+j, (T+j), P+(T+j)

直接短语: P, T, j

句柄: P

3. (12 分)

(1) (4 分) $E \rightarrow aTFb$

(2) (4 分)

$T \rightarrow bT'$

$T' \rightarrow baT' | \epsilon$

$F \rightarrow c$

$FIRST(E) = \{a\}$,

$FIRST(T) = \{b\}$,

$FIRST(T') = \{b, \epsilon\}$

$FIRST(F) = \{c\}$

$FOLLOW(E) = \{\#\}$

$FOLLOW(T) = \{c\}$

$FOLLOW(T') = \{c\}$

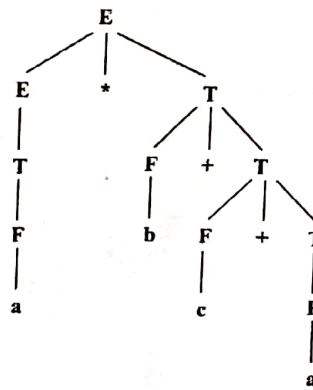
$FOLLOW(F) = \{b\}$

(3) (4 分)

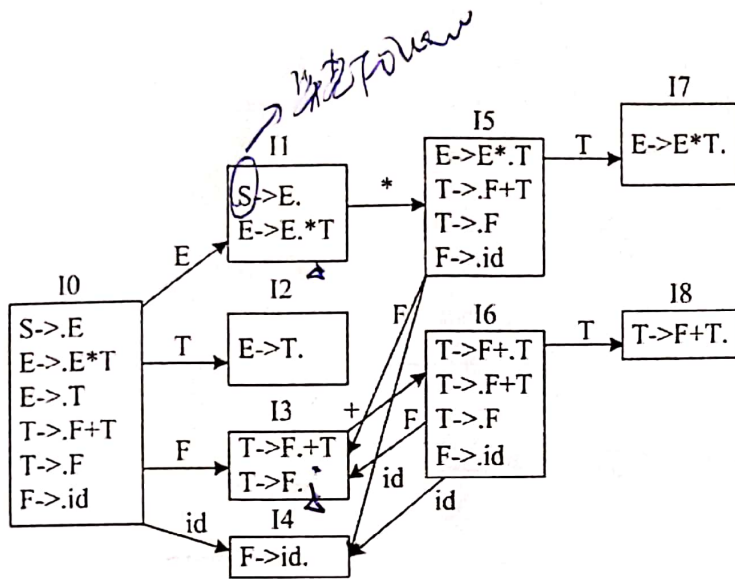
	a	b	c	#
E	aTFb			
T		bT'		
T'		baT'	ϵ	
F			c	

4. (12 分)

(1) (4 分)



(2) (4 分)



(3) (4 分)

$\text{FOLLOW}(S) = \{\#\}$, $\text{FOLLOW}(T) = \{\#, *\}$

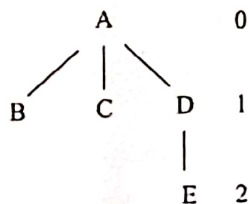
是 SLR(1) 文法, 对于 11: $* \notin \text{Follow}(S)$, 对于 13: $+ \notin \text{Follow}(T)$

5. (7 分)

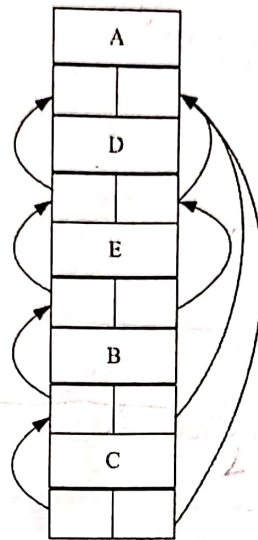
- (1) if a < c goto 5
- (2) goto 3
- (3) if b > d goto 5
- (4) goto ?
- (5) if a > 3 goto 7
- (6) goto ?
- (7) if b < 4 goto 9
- (8) goto ?
- (9) t1 := itr a
- (10) t2 := t1 + b
- (11) a := rti t2
- (12) goto 5

6. (8 分)

(1) (4 分)



(2) (4 分)



西安电子科技大学

考试时间_____分钟

试 题

题号	一	二	三	四	五	六	七	八	九	十	总分
分数											

1. 考试形式：闭（开）卷；2. 本试卷共 3 大题，满分 100 分。

班级_____学号_____姓名_____任课教师_____

一、填空题（20 分）

1.1 Lex 和 Yacc 是用于生成_____和_____的工具。

1.2 描述含 010 的 01 串的正规式是 $(0|1)^*0|0(0|1)^*$ 。

1.3 识别 3 型文法描述语言的自动机是 有限自动机。

1.4 在自上而下的语法分析中，需要消除左递归以避免无限循环，需要提取左因子以避免虚假匹配。

1.5 CFG 的定义包含有非终结符集合、终结符集合、和开始符号。

1.6 自下而上的分析中，需对输入序列进行规范规约，反复用产生式的左部替换句型中的产生式，最终得到开始符号。

1.7 文法产生二义性的原因是缺少对文法符号左递归和左因子的规定。

1.8 语义规则的两种表示方式是语义树和翻译方案。

1.9 4*4 二维数组 A 每维下界均为 1，每个元素占 1 个单元，若数组 A 的首地址为 a，并且以行为主存储，则元素 A[3, 2] 的地址是_____。

1.10 综合属性在分析树上的计算次序与自下而上语法分析形成分析树的过程一致。

1.11 LR 分析的语法制导翻译将语义规则放在产生式右部的右部，LR 分析器在执行移进动作时执行语义动作。

1.12 赋值语句 $x = a + (b * c)$ 的后缀式是 $x a b c + * =$ 。

1.13 CFG 无法描述语法中的变量声明与引用，可在语义规则中通过对符号表的插入、查找等操作实现。

1.14 参数传递的方法有值调用、引用调用、名字恢复。

换名调用。

1.15 动态分配策略有 栈分配 和堆分配策略。

二、简答题 (30 分)

2.1 (6 分) 解释什么是编译器的扫描遍数。

2.2 (6 分) 简述 DFA 与 NFA 有何区别。

2.3 (6 分) 解释什么是 LR(0) 项目集中的移进/归约冲突, 简单说明 SLR(1) 方法如何解决移进/归约冲突。

2.4 (6 分) 生成中间代码的目的是什么? 中间代码的特性是什么? 便于优化

2.5 (6 分) 控制栈中的活动记录保存的信息有什么? 具体内容有什么? 便于回溯

三、计算题 (50 分)

3.1 (12 分) 设有正规式 $r=a(a|b)^*b$, 试给出:

(1) (4 分) 识别该正规集的 NFA;

(2) (4 分) 识别该正规集的 DFA (要有计算过程);

(3) (4 分) 最小化的 DFA。

3.2 (12 分) 已知文法 G 如下: $S \rightarrow aABb$ $A \rightarrow Abc|d$ $B \rightarrow e$

(1) (4 分) 改写 G 为等价的 LL(1) 文法;

(2) (4 分) 求每个非终结符的 FIRST 集合和 FOLLOW 集合;

(3) (4 分) 构造预测分析表。

3.3 (14 分) 设有上下文无关文法 G 及其语法制导翻译如下 (注: G 中终结符 id 仅由单个英文字母组成, 如 a, b 等):

$E \rightarrow E1+T$ {E.place=newtemp; emit(+, E1.place, T.place, E.place;}

| T {E.place=T.place;}

$T \rightarrow T1*F$ {T.place=newtemp; emit(*, T1.place, F.place, T.place;}

| F {T.place=F.place;}

$F \rightarrow (E)$ {F.place=E.place;}

| id {F.place=id.name;}

(1) (3 分) 画出句子 $a+b*c+d$ 的分析树;

(2) (3 分) 写出当 $a=1, b=2, c=3, d=4$ 时的计算结果;

(3) (6 分) 将文法 G 简化为: $E \rightarrow E+T | T, T \rightarrow T*F | F, F \rightarrow id$, 给

出其识别活前缀的 DFA。

(4) (2 分) 该文法是 SLR(1) 文法吗? 为什么?

3.4 (12 分) 有一过程 A 如下所示。采用静态作用域、最近嵌套原则, 设 A 是第 0 层的过程。

```
procedure A is
  procedure B is x : integer; begin ..... end B;
  procedure C is x : integer; begin ...end C;
  procedure D is
    x : integer;
    procedure E is y: integer; begin .....end E;
  begin ..... end D;
begin .....end A;
```

(1) (4 分) 给出反映过程嵌套层次的嵌套树, 指出各过程的嵌套层次;

(2) (4 分) 给出正确反映作用域信息的符号表;

(3) (4 分) 若程序运行控制流是 A—B—D—E—B, 给出控制栈中的内容和控制链与访问链的指向。

一、填空题 (20 分)

- 1.1 词法分析器 语法分析器
- 1.2 $(0|1)^*010(0|1)^*$
- 1.3 有限自动机
- 1.4 死循环
- 1.5 产生式集合
- 1.6 句柄
- 1.7 优先级 结合性
- 1.8 语法制导定义 翻译方案
- 1.9 $a+9$
- 1.10 自下而上
- 1.11 最右边 规约
- 1.12 $xabc^{*}+=$
- 1.13 符号表
- 1.14 引用调用、复写-恢复
- 1.15 栈分配策略

二、简答题 (30 分)

2.1 (6 分) 编译器工作的每个阶段对以某种形式表示的完整程序进行一遍分析, 每个阶段将程序完整分析一遍的工作模式称为一遍扫描。

2.2 (6 分) DFA 与 NFA 的区别表现为两个方面: 一是 NFA 允许出现 ϵ 边。另一方面, DFA 的映象 f 是从状态集 S 和字母表 Σ 到 S 即 $S \times \Sigma \rightarrow S$, 而 NFA 的映象 f 是从状态集 S 和字母表 $\Sigma \cup \{\epsilon\}$ 到 S 幂集, 即, $S \times \Sigma \cup \{\epsilon\} \rightarrow 2^S$, 映象将产生一个状态集合 (可能为空集), 而不是单个状态。

2.3 (6 分) 移进/归约冲突是指在一个 LR(0) 项目集中同时存在可移进项 $A \rightarrow \beta_1 \beta_2$ 和可归约项 $B \rightarrow \beta_1$ 。对于移进/归约冲突, 计算非终结符 B 的 FOLLOW(B) 集合并看是否 $\text{FOLLOW}(B) \cap \text{FIRST}(\beta_2) = \Phi$, 若为 Φ 则可解决否则不可解决。

2.4 (6分) 中间代码的目的是便于编译器的开发移植和代码的优化。要求特性如下:

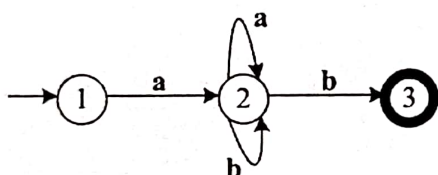
- ① 便于语法制导翻译;
- ② 与机器指令的结构相近, 又与具体机器无关。

2.5 (6分) 活动记录保存的有控制信息和访问信息, 具体内容有: 参数与返回值、控制链(可选)、访问链(可选)、调用时需要保存的机器状态、过程内部声明的数据、临时变量。

三、计算题 (50 分)

3.1 (12 分)

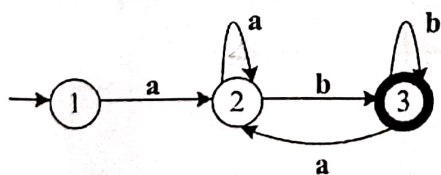
(1) (4 分)



(2) (4 分)

	a	b
{1}	{2}	
{2}	{2}	{2,3}
{2,3}	{2}	{2,3}

(3) (4 分)



3.2 (12 分)

(1) (4 分) 改写后的文法 $G'[S]: S \rightarrow aABb \quad A \rightarrow dA' \quad A' \rightarrow bcA' \mid \epsilon \quad B \rightarrow c$

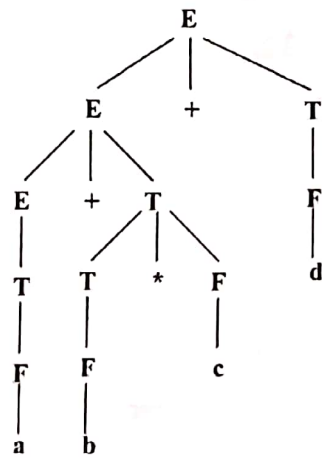
- (2) (4 分)
- | | |
|-------------------------------|------------------|
| FIRST(S) = {a}, | FOLLOW(S) = {#} |
| FIRST(A) = {d}, | FOLLOW(A) = {c} |
| FIRST(A') = {b, ϵ }, | FOLLOW(A') = {c} |
| FIRST(B) = {c}, | FOLLOW(B) = {b} |

(3) (1 分) 预测分析表如表所示:

	a	b	c	d	e	#
S	aABb					
A				dA'		
A'		bcA'			ϵ	
B					e	

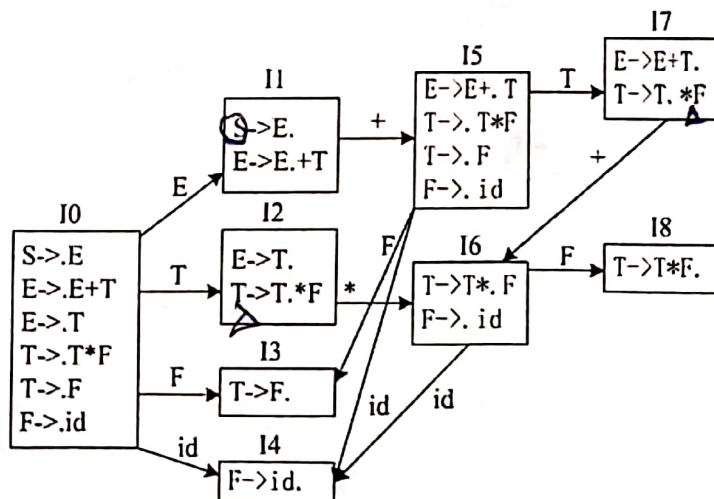
3.3 (14 分)

(1) (3 分)



(2) (3 分) $1+2*3+4=1+6+4=11$

(3) (6 分)



(4) (2 分)

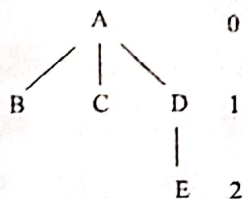
$\text{FIRST}(F)=\text{FIRST}(T)=\text{FIRST}(E)=\{\text{id}\}$

$FOLLOW(E) = \{ \#, + \}$, $FOLLOW(T) = \{ \#, +, * \}$, $FOLLOW(F) = \{ \#, +, * \}$
 是SLR(1)文法.

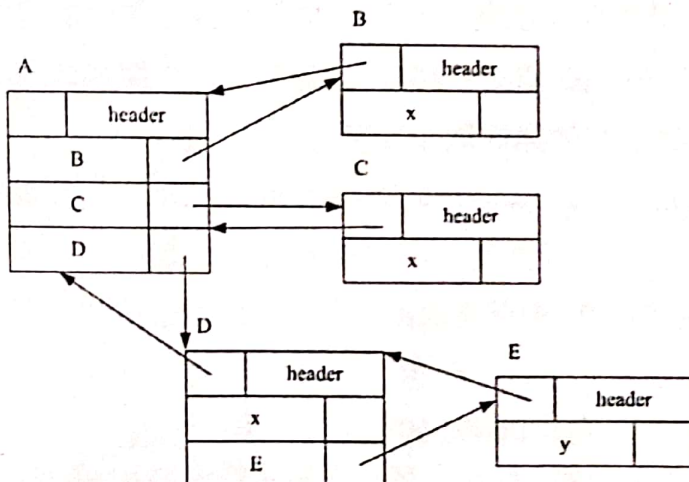
对于11: $+ \in Follow(S)$ 对于12: $* \notin Follow(E)$ 对于17: $* \notin Follow(E)$

3.4 (12分)

(1) (4分)



(2) (4分)



(3) (4分)

