

# 编译原理复习题

## 一、是非题

1. 计算机高级语言翻译成低级语言只有解释一种方式。(×)
3. 每个文法都能改写为 LL(1) 文法。(×)
4. 算符优先关系表不一定存在对应的优先函数。(√)
5. LR 分析方法是自顶向下语法分析方法。(×)
6. “用高级语言书写的源程序都必须通过编译,产生目标代码后才能投入运行”这种说法。(×)
7. 一个句型的句柄一定是文法某产生式的右部。(√)
8. 仅考虑一个基本块,不能确定一个赋值是否真是无用的。(√)
9. 在中间代码优化中循环上的优化主要有不变表达式外提和削减运算强度。(×)
10. 对于数据空间的存贮分配, FORTRAN 采用动态贮存分配策略。(×)
11. 甲机上的某编译程序在乙机上能直接使用的必要条件是甲机和乙机的操作系统功能完全相同。(×)
12. 递归下降分析法是自顶向下分析方法。(√)
13. 产生式是用于定义词法成分的一种书写规则。(×)
14. 在 SLR(1)分析法的名称中, S 的含义是简单的。(√)
15. 综合属性是用于“自上而下”传递信息。(×)
16. 符号表中的信息栏中登记了每个名字的属性和特征等有关信息,如类型、种属、所占单元大小、地址等等。(×)
17. 程序语言的语言处理程序是一种应用软件。(×)
18. 解释程序适用于 COBOL 和 FORTRAN 语言。(×)
19. 一个 LL(1)文法一定是无二义的。(√)
20. 正规文法产生的语言都可以用上下文无关文法来描述。(√)
21. 一张转换图只包含有限个状态,其中有一个被认为是初态,最多只有一个终态。(×)
22. 目标代码生成时,应考虑如何充分利用计算机的寄存器的问題。(√)
22. 逆波兰法表示的表达式亦称后缀式。(√)
23. 如果一个文法存在某个句子对应两棵不同的语法树,则称这个文法是二义的。(√)
24. 数组元素的地址计算与数组的存储方式有关。(√)
25. 算符优先关系表不一定存在对应的优先函数。(×)
26. 编译程序是对高级语言程序的解释执行。(×)
27. 一个有限状态自动机中,有且仅有一个唯一的终态。(×)
28. 一个算符优先文法可能不存在算符优先函数与之对应。(√)
29. 语法分析时必须先消除文法中的左递归。(×)
30. LR 分析法在自左至右扫描输入串时就能发现错误,但不能准确地指出出错地点。(√)
31. 逆波兰表示法表示表达式时无须使用括号。(√)
32. 静态数组的存储空间可以在编译时确定。(√)
33. 进行代码优化时应着重考虑循环的代码优化,这对提高目标代码的效率将起更大作用。(√)
34. 两个正规集相等的必要条件是他们对应的正规式等价。(√)
35. 一个语义子程序描述了一个文法所对应的翻译工作。(×)
36. 设  $r$  和  $s$  分别是正规式,则有  $L(r|s)=L(r)L(s)$ 。(×)
37. 确定的自动机以及不确定的自动机都能正确地识别正规集。(√)
38. 词法分析作为单独的一遍来处理较好。(×)
39. 构造 LR 分析器的任务就是产生 LR 分析表。(√)
40. 规范归约和规范推导是互逆的两个过程。(√)
41. 同心集的合并有可能产生新的“移进”/“归约”冲突。(×)
42. LR 分析技术无法适用二义文法。(×)
43. 树形表示和四元式不便于优化,而三元式和间接三元式则便于优化。(×)

44. 程序中的表达式语句在语义翻译时不需要回填技术。(√)  
 45. 对中间代码的优化依赖于具体的计算机。(×)  
 46. 若一个句型中出现了某产生式的右部, 则此右部一定是该句型的句柄。(×)  
 47. 在程序中标识符的出现仅为使用性的。(×)  
 48. 削减运算强度破坏了临时变量在一基本块内仅被定义一次的特性。(×)  
 49. 编译程序与具体的机器有关, 与具体的语言无关。(×)

二、选择题(请在前括号内选择最确切的一项作为答案划一个勾, 多划按错论)

1. 一个编译程序中, 不仅包含词法分析, ( A ), 中间代码生成, 代码优化, 目标代码生成等五个部分。  
 A. 语法分析 B. 文法分析 C. 语言分析 D. 解释分析
2. 语法分析器则可以发现源程序中的( D )。  
 A. 语义错误 B. 语法和语义错误 C. 错误并校正 D. 语法错误
3. 解释程序处理语言时, 大多数采用的是( B )方法。  
 A. 源程序命令被逐个直接解释执行  
 B. 先将源程序转化为中间代码, 再解释执行  
 C. 先将源程序解释转化为目标程序, 再执行  
 D. 以上方法都可以
4. 编译程序是一种( B )。  
 A. 汇编程序 B. 翻译程序 C. 解释程序 D. 目标程序
5. 文法分为四种类型, 即 0 型、1 型、2 型、3 型。其中 3 型文法是( B )。  
 A. 短语文法 B. 正则文法 C. 上下文有关文法 D. 上下文无关文法
6. 通常一个编译程序中, 不仅包含词法分析, 语法分析, 中间代码生成, 代码优化, 目标代码生成等五个部分, 还应包括( C )。  
 A. 模拟执行器 B. 解释器 C. 表格处理和出错处理 D. 符号执行器
7. 一个句型中的最左( B )称为该句型的句柄。  
 A. 短语 B. 简单短语 C. 素短语 D. 终结符号
8. 文法  $G[E]$  :  
 $E \rightarrow T \mid E + T$   
 $T \rightarrow F \mid T * F$   
 $F \rightarrow a \mid ( E )$   
 该文法句型  $E + F * (E + T)$  的简单短语是下列符号串中的( B )。  
 ①  $( E + T )$  ②  $E + T$  ③  $F$  ④  $F * (E + T)$   
 A. ① 和 ③ B. ② 和 ③ C. ③ 和 ④ D. ③
9. 词法分析器用于识别( C )。  
 A. 句子 B. 句型 C. 单词 D. 产生式
10. 在自底向上的语法分析方法中, 分析的关键是( A )。  
 A. 寻找句柄 B. 寻找句型 C. 消除递归 D. 选择候选式
11. 文法  $G$  产生的( D )的全体是该文法描述的语言。  
 A. 句型 B. 终结符集 C. 非终结符集 D. 句子
12. 若文法  $G$  定义的语言是无限集, 则文法必然是( A )。  
 A. 递归的 B. 前后文无关的 C. 二义性的 D. 无二义性的
13. 四种形式语言文法中, 1 型文法又称为( C )文法。  
 A. 短语结构文法 B. 前后文无关文法 C. 前后文有关文法 D. 正规文法
14. 一个文法所描述的语言是( A )。  
 A. 唯一的 B. 不唯一的 C. 可能唯一, 好可能不唯一 D. 都不对
15. ( B )和代码优化部分不是每个编译程序都必需的。  
 A. 语法分析 B. 中间代码生成 C. 词法分析 D. 目标代码生成
16. ( B )是两类程序语言处理程序。  
 A. 高级语言程序和低级语言程序 B. 解释程序和编译程序

- C. 编译程序和操作系统                      D. 系统程序和应用程序
17. 数组的内情向量中肯定不含有数组的(D)的信息。  
A. 维数              B. 类型              C. 维上下界              D. 各维的界差
18. 一个上下文无关文法 G 包括四个组成部分, 它们是: 一组非终结符号, 一组终结符号, 一个开始符号, 以及一组(D)。  
A. 句子              B. 句型              C. 单词              D. 产生式
19. 文法分为四种类型, 即 0 型、1 型、2 型、3 型。其中 2 型文法是(D)。  
A. 短语文法              B. 正则文法              C. 上下文有关文法              D. 上下文无关文法
20. 文法 G 所描述的语言是(C)的集合。  
A. 文法 G 的字母表 V 中所有符号组成的符号串              B. 文法 G 的字母表 V 的闭包  $V^*$  中的所有符号串  
C. 由文法的开始符号推出的所有终极符号              D. 由文法的开始符号推出的所有符号串
21. 词法分析器用于识别(C)。  
A. 字符串              B. 语句              C. 单词              D. 标识符
22. 文法分为四种类型, 即 0 型、1 型、2 型、3 型。其中 0 型文法是(A)。  
A. 短语文法              B. 正则文法              C. 上下文有关文法              D. 上下文无关文法
24. (A)是一种典型的解释型语言。  
A. BASIC              B. C              C. FORTRAN              D. PASCAL
25. 与编译系统相比, 解释系统(D)。  
A. 比较简单, 可移植性好, 执行速度快              B. 比较复杂, 可移植性好, 执行速度快  
C. 比较简单, 可移植性差, 执行速度慢              D. 比较简单, 可移植性好, 执行速度慢
26. 用高级语言编写的程序经编译后产生的程序叫(B)。  
A. 源程序              B. 目标程序              C. 连接程序              D. 解释程序
27. 词法分析器用于识别(A)。  
A. 字符串              B. 语句              C. 单词              D. 标识符
28. 编写一个计算机高级语言的源程序后, 到正式上机运行之前, 一般要经过(B)这几步:  
(1) 编辑 (2) 编译 (3) 连接 (4) 运行  
A. (1)(2)(3)(4)              B. (1)(2)(3)              C. (1)(3)              D. (1)(4)
29. 把汇编语言程序翻译成机器可执行的目标程序的工作是由(B)完成的。  
A. 编译器              B. 汇编器              C. 解释器              D. 预处理器
31. 词法分析器的输出结果是(C)。  
A. 单词的种别编码              B. 单词在符号表中的位置              C. 单词的种别编码和自身值              D. 单词自身值
32. 正规式 M1 和 M2 等价是指(C)。  
A. M1 和 M2 的状态数相等              B. M1 和 M2 的有向边条数相等  
C. M1 和 M2 所识别的语言集相等              D. M1 和 M2 状态数和有向边条数相等
33. 文法 G:  $S \rightarrow xSx|y$  所识别的语言是(C)。  
A.  $xyx$               B.  $(xyx)^*$               C.  $x^n y x^n (n \geq 0)$               D.  $x^* y x^*$
34. 如果文法 G 是无二义的, 则它的任何句子  $\alpha$  (A)。  
A. 最左推导和最右推导对应的语法树必定相同              B. 最左推导和最右推导对应的语法树可能不同  
C. 最左推导和最右推导必定相同              D. 可能存在两个不同的最左推导, 但它们对应的语法树相同
35. 构造编译程序应掌握(D)。  
A. 源程序              B. 目标语言              C. 编译方法              D. 以上三项都是
36. 四元式之间的联系是通过(B)实现的。  
A. 指示器              B. 临时变量              C. 符号表              D. 程序变量
37. 表达式  $(\neg A \vee B) \wedge (C \vee D)$  的逆波兰表示为(B)。  
A.  $\neg AB \vee \wedge CD \vee$               B.  $A \neg B \vee CD \vee \wedge$               C.  $AB \vee \neg CD \vee \wedge$               D.  $A \neg B \vee \wedge CD \vee$
38. 优化可生成(D)的目标代码。  
A. 运行时间较短              B. 占用存储空间较小

- C. 运行时间短但占用内存空间大      D. 运行时间短且占用存储空间小
39. 下列(C )优化方法不是针对循环优化进行的。
- A. 强度削弱      B. 删除归纳变量      C. 删除多余运算      D. 代码外提
40. 编译程序使用(B)区别标识符的作用域。
- A. 说明标识符的过程或函数名      B. 说明标识符的过程或函数的静态层次
- C. 说明标识符的过程或函数的动态层次      D. 标识符的行号
41. 编译程序绝大多数时间花在(D)上。
- A. 出错处理      B. 词法分析      C. 目标代码生成      D. 表格管理
42. 编译程序是对(D)。
- A. 汇编程序的翻译      B. 高级语言程序的解释执行      C. 机器语言的执行      D. 高级语言的翻译
43. 采用自上而下分析, 必须(C)。
- A. 消除左递归      B. 消除右递归      C. 消除回溯      D. 提取公共左因子
44. 在规范归约中, 用(B)来刻画可归约串。
- A. 直接短语      B. 句柄      C. 最左素短语      D. 素短语
45. 若  $a$  为终结符, 则  $A \rightarrow \alpha \cdot a\beta$  为(B)项目。
- A. 归约      B. 移进      C. 接受      D. 待约
46. 间接三元式表示法的优点为(A)。
- A. 采用间接码表, 便于优化处理      B. 节省存储空间, 不便于表的修改
- C. 便于优化处理, 节省存储空间      D. 节省存储空间, 不便于优化处理
47. 基本块内的优化为(B)。
- A. 代码外提, 删除归纳变量      B. 删除多余运算, 删除无用赋值
- C. 强度削弱, 代码外提      D. 循环展开, 循环合并
48. 在目标代码生成阶段, 符号表用(D)。
- A. 目标代码生成      B. 语义检查      C. 语法检查      D. 地址分配
49. 若项目集  $I_k$  含有  $A \rightarrow \alpha \cdot$ , 则在状态  $k$  时, 仅当面临的输入符号  $a \in FOLLOW(A)$  时, 才采取“ $A \rightarrow \alpha \cdot$ ”动作的一定是(D)。
- A. LALR 文法      B. LR(0)文法      C. LR(1)文法      D. SLR(1)文法
50. 堆式动态分配申请和释放存储空间遵守(D)原则。
- A. 先请先放      B. 先请后放      C. 后请先放      D. 任意

### 三、填空题

- 编译程序的工作过程一般可以划分为词法分析, 语法分析, 语义分析, 中间代码生成, 代码优化等几个基本阶段, 同时还会伴有\_\_表格处理\_\_和 \_\_出错处理\_\_。
- 编译方式与解释方式的根本区别在于\_\_是否生成目标代码\_\_。
- 产生式是用于定义\_\_语法成分\_\_的一种书写规则。
- 设  $G$  是一个给定的文法,  $S$  是文法的开始符号, 如果  $S \rightarrow x$  (其中  $x \in VT^*$ ), 则称  $x$  是文法的一个\_\_句子\_\_。
- 自顶向下的语法分析方法的基本思想是: 从文法的\_\_开始符号\_\_开始, 根据给定的输入串并按照文法的产生式一步一步的向下进行\_\_直接推导\_\_, 试图推导出文法的\_\_句子\_\_, 使之与给定的输入串\_\_匹配\_\_。
- 常用的参数传递方式有\_\_传地址\_\_, 传值和传名。
- 一个句型中的最左简单短语称为该句型的\_\_句柄\_\_。
- 对于文法的每个产生式都配备了一组属性的计算规则, 称为 \_\_语义规则\_\_。
- 一个典型的编译程序中, 不仅包括\_\_词法分析\_\_、\_\_语法分析\_\_、\_\_中间代码生成\_\_、代码优化、目标代码生成等五个部分, 还应包括表格处理和出错处理。
- 从功能上说, 程序语言的语句大体可分为\_\_执行性\_\_语句和\_\_说明性\_\_语句两大类。
- 扫描器的任务是从\_\_源程序\_\_中识别出一个个\_\_单词符号\_\_。
- 产生式是用于定义\_\_语法范畴\_\_的一种书写规则。
- 语法分析是依据语言的\_\_语法\_\_规则进行的, 中间代码产生是依据语言的\_\_语义\_\_规进行的。
- 语法分析器的输入是\_\_单词符号串\_\_, 其输出是\_\_语法单位\_\_。
- 一个名字的属性包括\_\_类型\_\_和\_\_作用域\_\_。

16. 逆波兰式  $ab+c+d*e$  所表达的表达式为  $(a+b+c)*d-e$ 。
17. 语法分析最常用的两类方法是自上而下和自下而上分析法。
18. 计算机执行用高级语言编写的程序主要有两种途径：解释和编译。
19. 扫描器是词法分析器，它接受输入的源程序，对源程序进行词法分析并识别出一个个单词符号，其输出结果是单词符号，供语法分析器使用。
20. 自上而下分析法采用移进、归约、错误处理、接受等四种操作。
21. 一个 LR 分析器包括两部分：一个总控程序和一张分析表。
22. 后缀式  $abc-/$  所代表的表达式是  $a/(b-c)$ 。
23. 局部优化是在基本块范围内进行的一种优化。
24. 词法分析基于正则文法进行，即识别的单词是该类文法的句子。
25. 语法分析基于上下文无关文法进行，即识别的是该类文法的句子。语法分析的有效工具是语法树。
26. 分析句型时，应用算符优先分析技术时，每步被直接归约的是最左素短语，而应用 LR 分析技术时，每步被直接归约的是句柄。
27. 语义分析阶段所生成的与源程序等价的中间表示形式可以有逆波兰、四元式表示与三元式表示等。
28. 按 Chomsky 分类法，文法按照规则定义的形式进行分类。
29. 一个文法能用有穷多个规则描述无穷的符号串集合（语言）是因为文法中存在有递归定义的规则。

#### 四、简答题

1. 写一文法，使其语言是偶正整数的集合，要求：

- (1) 允许 0 打头；
- (2) 不允许 0 打头。

解：(1)  $G[S] = (\{S, P, D, N\}, \{0, 1, 2, \dots, 9\}, P, S)$

P:

$S \rightarrow PD|D$

$P \rightarrow NP|N$

$D \rightarrow 0|2|4|6|8$

$N \rightarrow 0|1|2|3|4|5|6|7|8|9$

(2)  $G[S] = (\{S, P, R, D, N, Q\}, \{0, 1, 2, \dots, 9\}, P, S)$

P:

$S \rightarrow PD|P0|D$

$P \rightarrow NR|N$

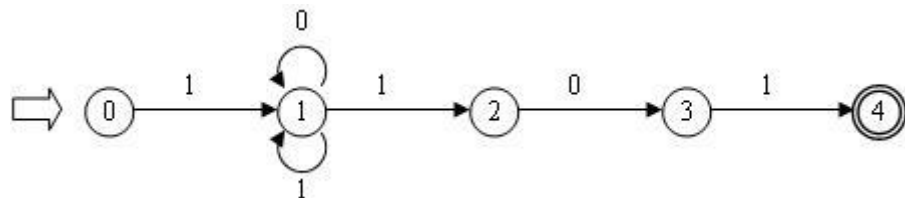
$R \rightarrow QR|Q$

$D \rightarrow 2|4|6|8$

$N \rightarrow 1|2|3|4|5|6|7|8|9$

$Q \rightarrow 0|1|2|3|4|5|6|7|8|9$

2. 构造正规式相应的 NFA :  $1(0|1)^*101$



解  $1(0|1)^*101$  对应的 NFA 为

3. 写出表达式  $(a+b*c)/(a+b)-d$  的逆波兰表示和三元式序列。

逆波兰表示:  $abc*+ab+/d-$

三元式序列:

- ①  $(*, b, c)$
- ②  $(+, a, ①)$
- ③  $(+, a, b)$

④ (/, ②, ③)

⑤ (一, ④, d)

4. 已知文法  $G[S]$  为:

$S \rightarrow dAB$

$A \rightarrow aA|a$

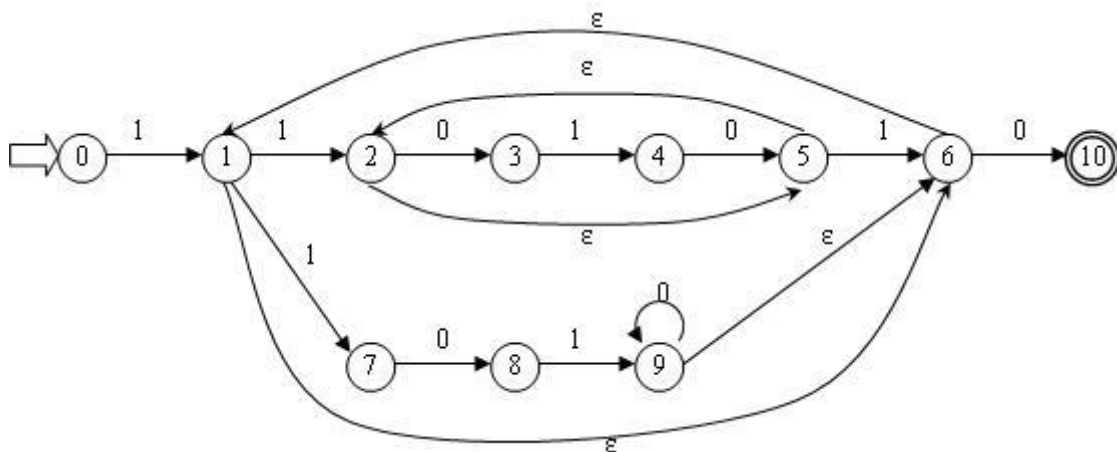
$B \rightarrow Bb|\epsilon$

$G[S]$  产生的语言是什么?

答:  $G[S]$  产生的语言是  $L(G[S]) = \{ da^n b^m \mid n \geq 1, m \geq 0 \}$ 。

5. 构造正规式相应的 DFA:  $1(1010^* | 1(010)^* 1)^* 0$ 。

解:  $1(1010^* | 1(010)^* 1)^* 0$  对应的 NFA 为:



6. 已知文法  $G(S)$

$S \rightarrow a | \wedge | (T)$

$T \rightarrow T, S | S$

写出句子  $((a, a), a)$  的规范归约过程及每一步的句柄。

解:

句型	归约规则	句柄
$((a, a), a)$	$S \rightarrow a$	$a$
$((S, a), a)$	$T \rightarrow S$	$S$
$((T, a), a)$	$S \rightarrow a$	$a$
$((T, S), a)$	$T \rightarrow T, S$	$T, S$
$((S), a)$	$T \rightarrow S$	$S$
$((T), a)$	$S \rightarrow S(T)$	$(T)$
$(S, a)$	$T \rightarrow S$	$S$
$(T, a)$	$S \rightarrow a$	$a$
$(T, S)$	$T \rightarrow T, S$	$T, S$
$(T)$	$S \rightarrow (T)$	$(T)$

$S$

7. 写一个文法, 使其语言是奇数集, 且每个奇数不以 0 开头。

解: 文法  $G(N)$ :

$N \rightarrow AB|B$

$A \rightarrow AC|D$

$B \rightarrow 1|3|5|7|9$

$D \rightarrow B|2|4|6|8$

$C \rightarrow 0|D$

8. 设文法  $G(S)$ :



$$S \rightarrow (L) | a S | a$$

$$L \rightarrow L, S | S$$

(1) 消除左递归和回溯;

(2) 计算每个非终结符的 FIRST 和 FOLLOW。

解: (1)

$$S \rightarrow (L) | a S'$$

$$S' \rightarrow S | \varepsilon$$

$$L \rightarrow S L'$$

$$L' \rightarrow S L' | \varepsilon$$

(2)

$$\text{FIRST}(S) = \{ (, a \} \quad \text{FOLLOW}(S) = \{ \#, , , ) \}$$

$$\text{FIRST}(S') = \{ , a, \varepsilon \} \quad \text{FOLLOW}(S') = \{ \#, , , ) \}$$

$$\text{FIRST}(L) = \{ (, a \} \quad \text{FOLLOW}(L) = \{ ) \}$$

$$\text{FIRST}(L') = \{ , , \varepsilon \} \quad \text{FOLLOW}(L') = \{ ) \}$$

9. 已知文法 G(E)

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

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

$$F \rightarrow (E) | i$$

(1) 给出句型  $(T * F + i)$  的最右推导;

(2) 给出句型  $(T * F + i)$  的短语、素短语。

解: (1) 最右推导:  $E \Rightarrow T \Rightarrow F \Rightarrow (E) \Rightarrow (E + T) \Rightarrow (E + F) \Rightarrow (E + i) \Rightarrow (T + i) \Rightarrow (T * F + i)$

(2) 短语:  $(T * F + i)$ ,  $T * F + i$ ,  $T * F$ ,  $i$

素短语:  $T * F, i$

10. While  $a > 0 \vee b < 0$  do

Begin

$X := X + 1;$

if  $a > 0$  then  $a := a - 1$

else  $b := b + 1$

End;

翻译成四元式序列。

解:

(1)  $(j >, a, 0, 5)$

(2)  $(j, -, -, 3)$

(3)  $(j <, b, 0, 5)$

(4)  $(j, -, -, 15)$

(5)  $(+, X, 1, T1)$

(6)  $(: =, T1, -, X)$

(7)  $(j \geq, a, 0, 9)$

(8)  $(j, -, -, 12)$

(9)  $(-, a, 1, T2)$

(10)  $(: =, T2, -, a)$

(11)  $(j, -, -, 1)$

(12)  $(+, b, 1, T3)$

(13)  $(: =, T3, -, b)$

(14)  $(j, -, -, 1)$

(15)

11. 写出下列表达式的三地址形式的中间表示。

(1)  $5 + 6 * (a + b);$

(2)for j:=1 to 10 do a[j + j]:=0。

答： (1)100: t1:=a+b

101: t2:=6\*t1

102: t3:=5+t2

(2)100: j:=1

101: if j>10 goto NEXT

102: i:=j+j

103: a[i]:=0

12. 设基本块 p 由如下语句构成：

T 0 :=3.14;

T 1 :=2\*T 0 ;

T 2 :=R+r;

A:=T 1 \*T 2 ;

B:=A;

T 3 :=2\*T 0 ;

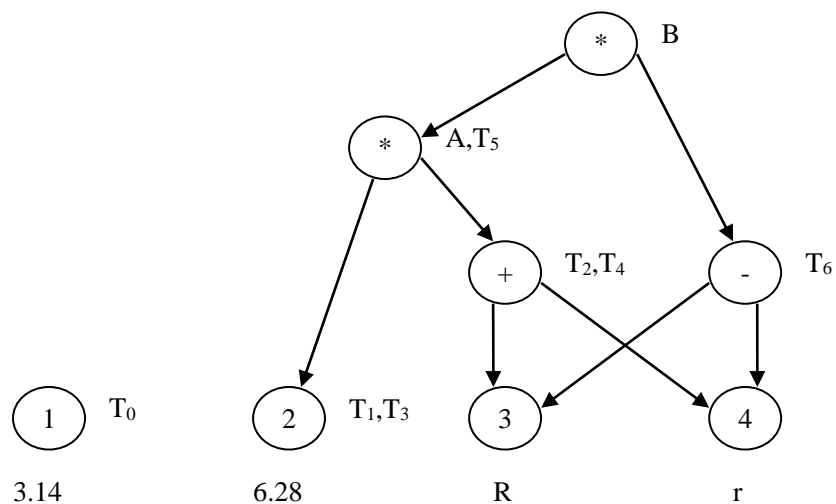
T 4 :=R+r;

T 5 :=T 3 \*T 4 ;

T 6 :=R-r ;

B:=T 5 \*T 6 ;

试给出基本块 p 的 DAG 。



解：基本块 p 的 DAG 图：

13. 写出表达式(a+b)/(a-b-(a+b\*c)的三元序列及四元序列。

解：(1) 三元式：

① (+, a, b)

② (-, a, b)

③ (/ , ①, ②)

④ (\*, b, c)

⑤ (+, a, ④)

⑥ (-, ③, ⑤)

(2) 四元式：

① (+, a, b, T1)

② (-, a, b, T2)

③ (/ , T1, T2, T3)

④ (\*, b, c, T4)

⑤ (+, a, T4, T5)



⑥ (一, T3, T5, T6)

14. 写一个文法使其语言为偶数集, 且每个偶数不以 0 开头。

解: 文法  $G(S)$ :

$S \rightarrow AB|B|A0$

$A \rightarrow AD|C$

$B \rightarrow 2|4|6|8$

$C \rightarrow 1|3|5|7|9|B$

$D \rightarrow 0|C$

15. 设文法  $G(S)$ :

$S \rightarrow S + aF|aF| + aF$

$F \rightarrow *aF|*a$

(1) 消除左递归和回溯;

(2) 构造相应的 FIRST 和 Follow 集合。

解: (1)

$S \rightarrow aFS'| + aFS'$

$S' \rightarrow + aFS'|\epsilon$

$F \rightarrow *aF'$

$F' \rightarrow F|\epsilon$

(2)

$FIRST(S) = \{a, +\}$        $FOLLOW(S) = \{\#\}$

$FIRST(S') = \{+, \epsilon\}$        $FOLLOW(S') = \{\#\}$

$FIRST(F) = \{*\}$        $FOLLOW(F) = \{+, \#\}$

$FIRST(F') = \{*, \epsilon\}$        $FOLLOW(+, \#)$

16. 简要说明语义分析的基本功能。

答: 语义分析的基本功能包括: 确定类型、类型检查、语义处理和某些静态语义检查。

17. 考虑文法  $G[S]$ :

$S \rightarrow (T) | a+S | a$

$T \rightarrow T,S | S$

消除文法的左递归及提取公共左因子。

解: 消除文法  $G[S]$  的左递归:

$S \rightarrow (T) | a+S | a$

$T \rightarrow ST'$

$T' \rightarrow ,ST'|\epsilon$

提取公共左因子:

$S \rightarrow (T) | aS'$

$S' \rightarrow +S|\epsilon$

$T \rightarrow ST'$

$T' \rightarrow ,ST'|\epsilon$

18. 试为表达式  $w+(a+b)*(c+d/(e-10)+8)$  写出相应的逆波兰表示。

解:  $w a b + c d e 10 - / + 8 + * +$

19. 按照三种基本控制结构文法将下面的语句翻译成四元式序列:

while ( $A < C \wedge B < D$ )

{

if ( $A \geq 1$ )  $C = C + 1$ ;

else while ( $A \leq D$ )

$A = A + 2$ ;

}。

解: 该语句的四元式序列如下(其中  $E_1$ 、 $E_2$  和  $E_3$  分别对应  $A < C \wedge B < D$ 、 $A \geq 1$  和  $A \leq D$ , 并且关系运算符优先级

高):

100 ( $j < A, C, 102$ )

101 ( $j, \_, \_, 113$ )

102 ( $j < B, D, 104$ )

103 ( $j, \_, \_, 113$ )

104 ( $j = A, 1, 106$ )

105 ( $j, \_, \_, 108$ )

106 (+, C, 1, C)

107 ( $j, \_, \_, 112$ )

108 ( $j \leq A, D, 110$ )

109 ( $j, \_, \_, 112$ )

110 (+, A, 2, A)

111 ( $j, \_, \_, 108$ )

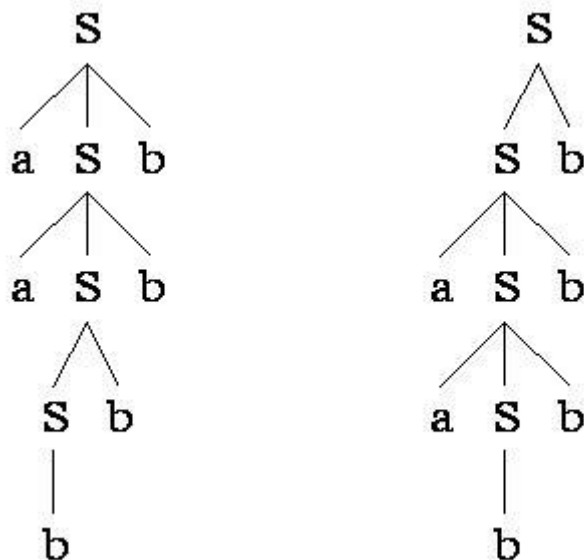
112 ( $j, \_, \_, 100$ )

113

20. 已知文法  $G[S]$  为  $S \rightarrow aSb|Sb|b$ ，试证明文法  $G[S]$  为二义文法。

证明:

由文法  $G[S]: S \rightarrow aSb|Sb|b$ ，对句子  $aabbbb$  对应的两棵语法树为:



因此，文法  $G[S]$  为二义文法。

21. 文法  $G[S]$  为:

$S \rightarrow Ac|aB$

$A \rightarrow ab$

$B \rightarrow bc$

写出  $L(G[S])$  的全部元素。

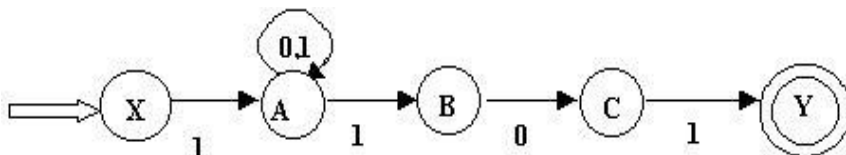
解:  $S \Rightarrow Ac \Rightarrow abc$

或  $S \Rightarrow aB \Rightarrow abc$

所以  $L(G[S]) = \{abc\}$

22. 构造正规式  $1(0|1)^*101$  相应的 DFA。

解: 先构造 NFA:



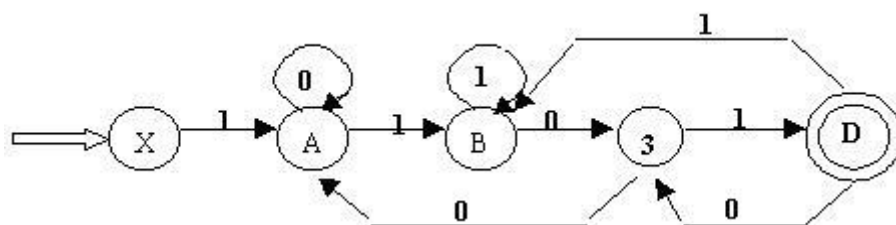
确定化:

	0	1
X		A
A	A	AB
AB	AC	AB
AC	A	ABY
ABY	AC	AB

重新命名, 令 AB 为 B、AC 为 C、ABY 为 D 得:

	0	1
X		A
A	A	B
B	C	B
C	A	D
D	C	B

所以, 可得 DFA 为:



23. 文法

$S \rightarrow a \mid (T)$

$T \rightarrow T, S \mid S$

对  $(a, (a, a))$  和  $((((a, a), ^{(a)}), a))$  的最左推导。

解: 对  $(a, (a, a))$  的最左推导为:

$S \Rightarrow (T) \Rightarrow (T, S) \Rightarrow (S, S) \Rightarrow (a, S)$

$\Rightarrow (a, (T)) \Rightarrow (a, (T, S)) \Rightarrow (a, (S, S))$

$\Rightarrow (a, (a, S)) \Rightarrow (a, (a, a))$

对  $((((a, a), ^{(a)}), a))$  的最左推导为:

$S \Rightarrow (T) \Rightarrow (T, S) \Rightarrow (S, S) \Rightarrow ((T), S)$

$\Rightarrow ((T, S), S) \Rightarrow ((T, S, S), S) \Rightarrow ((S, S, S), S)$

$\Rightarrow (((T), S, S), S) \Rightarrow (((T, S), S, S), S) \Rightarrow (((S, S), S, S), S)$

$\Rightarrow (((a, S), S, S), S) \Rightarrow (((a, a), S, S), S) \Rightarrow (((a, a), ^{(a)}), S)$

$\Rightarrow (((a, a), ^{(T)}), S) \Rightarrow (((a, a), ^{(S)}), S) \Rightarrow (((a, a), ^{(a)}), S)$

$\Rightarrow (((a, a), ^{(a)}), a)$

24. 文法:

$S \rightarrow MH \mid a$

$H \rightarrow LSo \mid \varepsilon$

$K \rightarrow dML \mid \varepsilon$

$L \rightarrow eHf$

$M \rightarrow K \mid bLM$

判断 G 是否为 LL(1) 文法, 如果是, 构造 LL(1) 分析表。

解: 各符号的 FIRST 集和 FOLLOW 集为:

	FIRST	FOLLOW
S	{a,d,b,ε,e}	{#,o}
M	{d,ε,b}	{e, #, o}
H	{ε,e}	{#,f,o}
L	{e}	{a,d,b,e,o,#}
K	{d,ε}	{e, #, o}

预测分析表为：

	a	o	d	e	f	b	#
S	->a	->MH	->MH	->MH		->MH	->MH
M		->K	->K	->K		->bLM	->K
H		->ε		->LS <sub>o</sub>	->ε		->ε
L				->eHf			
K		->ε	->dML	->ε			->ε

由于预测分析表中无多重入口，所以可判定文法是 LL(1)的。

25. 叙述由下列正规式描述的语言

(a)0(0|1)\*0

(b)((ε|0)1\*)\*

(c)(0|1)\*0(0|1)(0|1)

(d)0\*10\*10\*10\*

(e)(00|11)\*((01|10)(00|11)\*(01|10)(00|11)\*)\*

解：(a)以 0 开头、以 0 结尾的所有 0 和 1 的串。

(b)由 0 和 1 组成的串，包括空串。

(c)倒数第 3 个字符为 0，由 0 和 1 组成的串。

(d)含有 3 个 1 的所有 0 和 1 的串。

(e)由偶数个 0 和偶数个 1 构成的所有 0 和 1 的串。

26. 已知文法 G[S]：

$S \rightarrow (L) \mid a$

$L \rightarrow L, S \mid S$

为句子(a,(a,a))构造最左推导和最右推导。

解：句子(a,(a,a))的最左推导为：

$S \Rightarrow (L) \Rightarrow (L, S) \Rightarrow (S, S) \Rightarrow (a, S) \Rightarrow (a, (L)) \Rightarrow (a, (L, S)) \Rightarrow (a, (S, S)) \Rightarrow (a, (a, S)) \Rightarrow (a, (a, a))$

句子(a,(a,a))的最右推导为：

$S \Rightarrow (L) \Rightarrow (L, S) \Rightarrow (L, (L)) \Rightarrow (L, (L, S)) \Rightarrow (L, (L, a)) \Rightarrow (L, (S, a)) \Rightarrow (L, (a, a)) \Rightarrow (S, (a, a)) \Rightarrow (a, (a, a))$

五.计算题

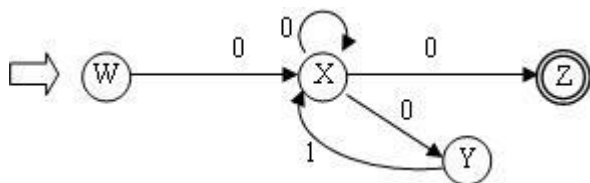
1. 构造下述文法 G[S] 的自动机：

$S \rightarrow A0$

$A \rightarrow A0 \mid S1 \mid 0$

该自动机是确定的吗？若不确定，则对它确定化。

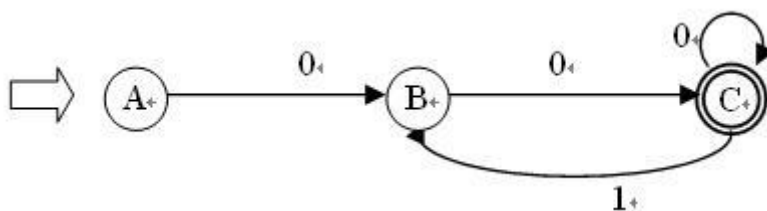
解：由于该文法的产生式  $S \rightarrow A0$ ， $A \rightarrow A0 \mid S1 \mid 0$  中没有字符集 VT 的输入，所以不是确定的自动机。要将其他确定化，必须先用代入法得到它对应的正规式。把  $S \rightarrow A0$  代入产生式  $A \rightarrow S1$  有： $A = A0 \mid A01 \mid 0 = A(0|01) \mid 0 = 0(0|01)^*$ 。代入  $S \rightarrow A0$  有该文法的正规式： $0(0|01)^*0$ ，所以，改写该文法为确定的自动机为：



由于状态 A 有 3 次输入 0 的重复输入，所以上图只是 NFA，下面将它确定化：

下表由子集法将 NFA 转换为 DFA:

I	$I_0 = \epsilon\text{-closure}(\text{MoveTo}(I, 0))$	$I_1 = \epsilon\text{-closure}(\text{MoveTo}(I, 1))$
A[W]	B[X]	
B[X]	C[X, Y, Z]	
C[X, Y, Z]	C[X, Y, Z]	B[X]



由上表可知 DFA 为：

2. 对下面的文法 G :

$E \rightarrow TE'$

$E' \rightarrow +E \mid \epsilon$

$T \rightarrow FT'$

$T' \rightarrow T \mid \epsilon$

$F \rightarrow PF'$

$F' \rightarrow *F' \mid \epsilon$

$P \rightarrow (E) \mid a \mid b \mid ^$

(1) 计算这个文法的每个非终结符的 FIRST 集和 FOLLOW 集。

(2) 证明这个方法是 LL(1) 的。

(3) 构造它的预测分析表。

解：(1) 计算这个文法的每个非终结符的 FIRST 集和 FOLLOW 集。

FIRST 集合有：

$\text{FIRST}(E) = \text{FIRST}(T) = \text{FIRST}(F) = \text{FIRST}(P) = \{(a, b, ^)\};$

$\text{FIRST}(E') = \{+, \epsilon\}$

$\text{FIRST}(T) = \text{FIRST}(F) = \text{FIRST}(P) = \{(a, b, ^)\};$

$\text{FIRST}(T') = \text{FIRST}(T) \cup \{\epsilon\} = \{(a, b, ^, \epsilon)\};$

$\text{FIRST}(F) = \text{FIRST}(P) = \{(a, b, ^)\};$

$\text{FIRST}(F') = \text{FIRST}(P) = \{*, \epsilon\};$

$\text{FIRST}(P) = \{(a, b, ^)\};$

FOLLOW 集合有：

$\text{FOLLOW}(E) = \{), \#\};$

$\text{FOLLOW}(E') = \text{FOLLOW}(E) = \{), \#\};$

$\text{FOLLOW}(T) = \text{FIRST}(E') \cup \text{FOLLOW}(E) = \{+, \epsilon, \#\};$  // 不包含  $\epsilon$

$\text{FOLLOW}(T') = \text{FOLLOW}(T) = \text{FIRST}(E') \cup \text{FOLLOW}(E) = \{+, \epsilon, \#\};$

$\text{FOLLOW}(F) = \text{FIRST}(T') \cup \text{FOLLOW}(T) = \{(a, b, ^, +, \epsilon, \#\};$  // 不包含  $\epsilon$

$\text{FOLLOW}(F') = \text{FOLLOW}(F) = \text{FIRST}(T') \cup \text{FOLLOW}(T) = \{(a, b, ^, +, \epsilon, \#\};$

$\text{FOLLOW}(P) = \text{FIRST}(F') \cup \text{FOLLOW}(F) = \{*, (a, b, ^, +, \epsilon, \#\};$  // 不包含  $\epsilon$

(2) 证明这个方法是 LL(1) 的。

各产生式的 SELECT 集合有:

SELECT( $E \rightarrow TE'$ )= $FIRST(T)=\{(,a,b,\wedge\}$ ;  
SELECT( $E' \rightarrow +E$ )= $\{+\}$ ;  
SELECT( $E' \rightarrow \epsilon$ )= $FOLLOW(E)=\{),\#\}$   
SELECT( $T \rightarrow FT'$ )= $FIRST(F)=\{(,a,b,\wedge\}$ ;  
SELECT( $T' \rightarrow T$ )= $FIRST(T)=\{(,a,b,\wedge\}$ ;  
SELECT( $T' \rightarrow \epsilon$ )= $FOLLOW(T)=\{+,\),\#\}$ ;  
SELECT( $F \rightarrow PF'$ )= $FIRST(P)=\{(,a,b,\wedge\}$ ;  
SELECT( $F' \rightarrow *F'$ )= $\{*\}$ ;  
SELECT( $F' \rightarrow \epsilon$ )= $FOLLOW(F')=\{(,a,b,\wedge,+,\),\#\}$ ;  
SELECT( $P \rightarrow (E)$ )= $\{($   
SELECT( $P \rightarrow a$ )= $\{a\}$   
SELECT( $P \rightarrow b$ )= $\{b\}$   
SELECT( $P \rightarrow \wedge$ )= $\{\wedge\}$

可见，相同左部产生式的 SELECT 集的交集均为空，所以文法  $G[E]$  是 LL(1) 文法。

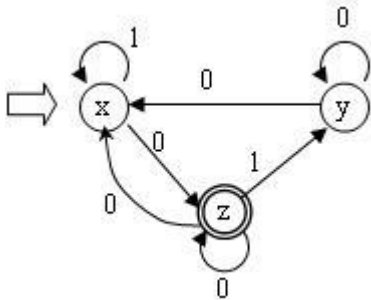
(3)构造它的预测分析表。

文法  $G[E]$  的预测分析表如下:

	+	*	(	)	a	b	^	#
E			$\rightarrow TE'$		$\rightarrow TE'$	$\rightarrow TE'$	$\rightarrow TE'$	
E'	$\rightarrow +E$			$\rightarrow \epsilon$				$\rightarrow \epsilon$
T			$\rightarrow FT'$		$\rightarrow FT'$	$\rightarrow FT'$	$\rightarrow FT'$	
T'	$\rightarrow \epsilon$		$\rightarrow T$	$\rightarrow \epsilon$	$\rightarrow T$	$\rightarrow T$	$\rightarrow T$	$\rightarrow \epsilon$
F			$\rightarrow PF'$		$\rightarrow PF'$	$\rightarrow PF'$	$\rightarrow PF'$	
F'	$\rightarrow \epsilon$	$\rightarrow *F'$	$\rightarrow \epsilon$	$\rightarrow \epsilon$	$\rightarrow \epsilon$	$\rightarrow \epsilon$	$\rightarrow \epsilon$	$\rightarrow \epsilon$
P			$\rightarrow (E)$		$\rightarrow a$	$\rightarrow b$	$\rightarrow \wedge$	

3. 已知  $NFA = (\{x,y,z\},\{0,1\},M,\{x\},\{z\})$ ，其中:

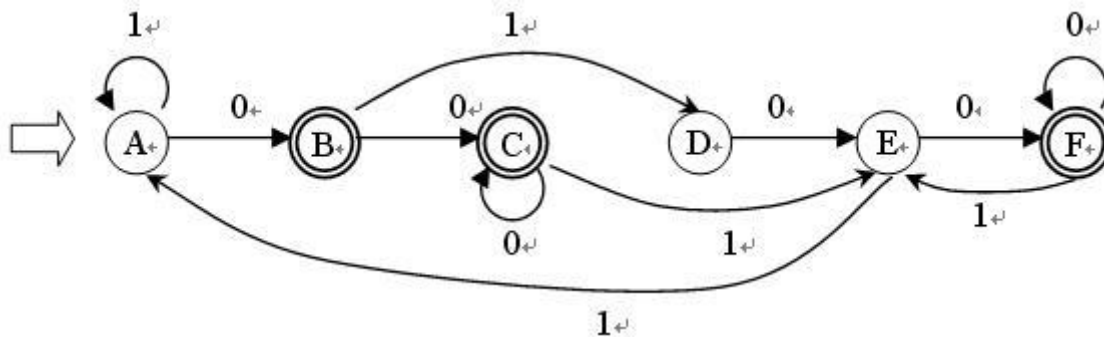
$M(x,0)=\{z\},M(y,0)=\{x,y\},M(z,0)=\{x,z\},M(x,1)=\{x\},M(y,1)=\varnothing,M(z,1)=\{y\}$ ，构造相应的 DFA 并最小化。



解: 根据题意有 NFA 图:

下表由子集法将 NFA 转换为 DFA :

I	$I_0 = \epsilon\text{-closure}(\text{MoveTo}(I, 0))$	$I_1 = \epsilon\text{-closure}(\text{MoveTo}(I, 1))$
A[x]	B[z]	A[x]
B[z]	C[x, z]	D[y]
C[x, z]	C[x, z]	E[x, y]
D[y]	E[x, y]	
E[x, y]	F[x, y, z]	A[x]
F[x, y, z]	F[x, y, z]	E[x, y]



下面将该 DFA 最小化:

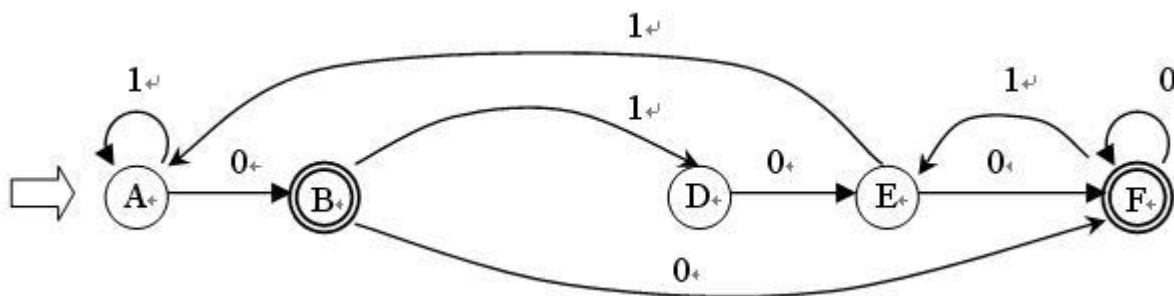
(1) 首先将它的状态集分成两个子集:  $P1=\{A,D,E\}, P2=\{B,C,F\}$

(2) 区分  $P2$ : 由于  $F(F,1)=F(C,1)=E, F(F,0)=F$  并且  $F(C,0)=C$ , 所以  $F, C$  等价。由于  $F(B,0)=F(C,0)=C, F(B,1)=D, F(C,1)=E$ , 而  $D, E$  不等价 (见下步), 从而  $B$  与  $C, F$  可以区分。有  $P21=\{C,F\}, P22=\{B\}$ 。

(3) 区分  $P1$ : 由于  $A, E$  输入 0 到终态, 而  $D$  输入 0 不到终态, 所以  $D$  与  $A, E$  可以区分, 有  $P11=\{A,E\}, P12=\{D\}$ 。

(4) 由于  $F(A,0)=B, F(E,0)=F$ , 而  $B, F$  不等价, 所以  $A, E$  可以区分。

(5) 综上所述, DFA 可以区分为  $P=\{\{A\}, \{B\}, \{D\}, \{E\}, \{C, F\}\}$ 。所以最小化的 DFA 如下:



4. 已知文法为:

$S \rightarrow a | \wedge | (T)$

$T \rightarrow T, S | S$

构造它的 LR(0)分析表。

解: 加入非终结符  $S'$ , 方法的增广文法为:

$S' \rightarrow S$

$S \rightarrow a$

$S \rightarrow \wedge$

$S \rightarrow (T)$

$T \rightarrow T, S$

$T \rightarrow S$  下面构造它的 LR(0)项目集规范族为:



状态 \ 当前符号	a	^	(	)	,	#	S	T
I <sub>0</sub> : S' → • S S → • a S → • ^ S → • (T)	I <sub>2</sub> : S → a •	I <sub>3</sub> : S → ^ •	I <sub>1</sub> : S → ( • T) T → • T, S T → • S S → • a S → • ^ S → • (T)				I <sub>11</sub> : S' → S •	
I <sub>1</sub>						acc		
I <sub>2</sub>								
I <sub>3</sub>								
I <sub>4</sub> : S → ( • T) T → • T, S T → • S S → • a S → • ^ S → • (T)	I <sub>2</sub>	I <sub>3</sub>	I <sub>1</sub>				I <sub>6</sub> : T → S •	I <sub>5</sub> : S → (T •) T → T •, S
I <sub>5</sub> : S → (T •) T → T •, S				I <sub>7</sub> : S → (T) •	I <sub>8</sub> : T → T, • S S → • a S → • ^ S → • (T)			
I <sub>6</sub>								
I <sub>7</sub> :								
I <sub>8</sub> : T → T, • S S → • a S → • ^ S → • (T)	I <sub>2</sub>	I <sub>3</sub>	I <sub>1</sub>				I <sub>9</sub> : T → T, S •	
I <sub>9</sub>								

从上表可看出,不存在移进-归约冲突以及归约归约冲突,该文法是 LR(0)文法。

从而有下面的 LR(0)分析表:

状态	ACTION						GOTO	
	a	^	(	)	,	#	S	T
0	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>				1	
1						acc		
2	r <sub>1</sub>	r <sub>1</sub>	r <sub>1</sub>	r <sub>1</sub>	r <sub>1</sub>	r <sub>1</sub>		
3	r <sub>2</sub>	r <sub>2</sub>	r <sub>2</sub>	r <sub>2</sub>	r <sub>2</sub>	r <sub>2</sub>		
4	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>				6	5
5				S <sub>7</sub>	S <sub>8</sub>			
6	r <sub>5</sub>	r <sub>5</sub>	r <sub>5</sub>	r <sub>5</sub>	r <sub>5</sub>	r <sub>5</sub>		
7	r <sub>3</sub>	r <sub>3</sub>	r <sub>3</sub>	r <sub>3</sub>	r <sub>3</sub>	r <sub>3</sub>		
8	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>				9	
9	r <sub>4</sub>	r <sub>4</sub>	r <sub>4</sub>	r <sub>4</sub>	r <sub>4</sub>	r <sub>4</sub>		

5. 已知文法

$A \rightarrow aAd|aAb|\epsilon$

判断该文法是否是 SLR(1) 文法,若是构造相应分析表,并对输入串 ab# 给出分析过程。

解: 增加一个非终结符 S/后, 产生原文法的增广文法有:

$S' \rightarrow A$

$A \rightarrow aAd|aAb|\epsilon$

下面构造它的 LR(0) 项目集规范族为：

状态 \ 当前符号	a	b	d	#	A
$I_0$ : $S' \rightarrow \cdot A$ $A \rightarrow \cdot aAd$ $A \rightarrow \cdot aAb$ $A \rightarrow \cdot$	$I_2$ : $A \rightarrow a \cdot Ad$ $A \rightarrow a \cdot Ab$ $A \rightarrow \cdot aAd$ $A \rightarrow \cdot aAb$ $A \rightarrow \cdot$				$I_1$ : $S' \rightarrow A \cdot$
$I_1$ : $S' \rightarrow A \cdot$				acc	
$I_2$ : $A \rightarrow a \cdot Ad$ $A \rightarrow a \cdot Ab$ $A \rightarrow \cdot aAd$ $A \rightarrow \cdot aAb$ $A \rightarrow \cdot$	$I_2$				$I_3$ : $A \rightarrow aA \cdot d$ $A \rightarrow aA \cdot b$
$I_3$ : $A \rightarrow aA \cdot d$ $A \rightarrow aA \cdot b$		$I_4$ : $A \rightarrow aAb \cdot$	$I_5$ : $A \rightarrow aAd \cdot$		
$I_4$ : $A \rightarrow aAb \cdot$					
$I_5$ : $A \rightarrow aAd \cdot$					

从上表可看出,状态  $I_0$  和  $I_2$  存在移进-归约冲突，该文法不是 LR(0)文法。对于  $I_0$  来说有： $FOLLOW(A) \cap \{a\} = \{b,d,\#\} \cap \{a\} = \Phi$ ，所以在  $I_0$  状态下面临输入符号为 a 时移进，为 b,d,# 时归约，为其他时报错。对于  $I_2$  来说也有与  $I_0$  完全相同的结论。这就是说，以上的移进-归约冲突是可以解决的，因此该文法是 SLR(1) 文法。

其 SLR(1)分析表为：

状态	ACTION				GOTO
	a	b	d	#	A
0	$S_2$	$r_1$	$r_2$	$r_3$	1
1				acc	
2	$S_2$	$r_1$	$r_2$	$r_3$	3
3		$S_4$	$S_5$		
4	$r_2$	$r_2$	$r_2$	$r_2$	
5	$r_1$	$r_1$	$r_1$	$r_1$	

对输入串 ab#给出分析过程为：

步骤	状态栈	符号栈	输入串	ACTION	GOTO
1	0	#	ab#	$S_2$	
2	02	#a	b#	$r_3$	3
3	023	#aA	b#	$S_4$	
4	0234	#aAb	#	$r_2$	1
5	01	#A	#	acc	

6. 已知文法  $G[S]$  为：

$S \rightarrow a \wedge (T)$

$T \rightarrow T, S | S$

(1) 计算  $G[S]$  的 FIRSTVT 和 LASTVT。

- (2) 构造  $G[S]$  的算符优先关系表并说明  $G[S]$  是否未算符优先文法。
- (3) 计算  $G[S]$  的优先函数。
- (4) 给出输入串  $(a,a)\#$  的算符优先分析过程。

解：（1）各符号的 FIRSTVT 和 LASTVT：

	FIRSTVT	LASTVT
S	a、^、(	a、^、)
T	,、a、^、(	,、a、^、)

（2）算符优先关系表：

	a	(	)	,	^	#
a			$\triangleright$	$\triangleright$		$\triangleright$
(	$\lessdot$	$\lessdot$	$\equiv$		$\lessdot$	
)			$\triangleright$	$\triangleright$		$\triangleright$
,	$\lessdot$	$\lessdot$	$\triangleright$	$\triangleright$	$\lessdot$	
^			$\triangleright$	$\triangleright$		$\triangleright$
#	$\lessdot$	$\lessdot$			$\lessdot$	

（3）对应的算符优先函数为：

	a	(	)	,	^	#
S	2	1	2	2	2	1
T	3	3	1	1	3	1

（4）句子 $(a,a)\#$ 分析过程如下：

步骤	栈	优先关系	当前符号	剩余输入串	移进或归约
1	#	$\# \lessdot ($	(	a,a)#	移进
2	$\# ($	$( \lessdot a$	a	,a)#	移进
3	$\# (a$	$a \triangleright ,$	,	a)#	归约
4	$\# (F$	$( \lessdot ,$	,	a)#	移进
5	$\# (F,$	$, \lessdot a$	A	)#	移进
6	$\# (F,a$	$A \triangleright )$	)	#	归约
7	$\# (F,F$	$, \triangleright )$	)	#	归约
8	$\# (F$	$( \equiv )$	)	#	移进
9	$\# (F)$	$) \triangleright \#$	#		归约
10	$\# F$	$\# \equiv \#$	#		接受