

# JLU 编译原理试题解答

## 目录

<a href="#">知识点总结</a>	1
<a href="#">试题样例</a>	4
<a href="#">2003 级编译原理试题(Done)</a>	5
<a href="#">2004 级编译原理试题</a>	8
<a href="#">2007 级编译原理试题 1</a>	11
<a href="#">2007 级编译原理试题 2(Done)</a>	14
<a href="#">2008 级编译原理试题</a>	17
<a href="#">2010 级编译原理试题</a>	21
<a href="#">2011 级编译原理试题</a>	25
<a href="#">2012 级编译原理试题(Done)</a>	30
<a href="#">2013 级编译原理试题 1(Done)</a>	37
<a href="#">2013 级编译原理试题 2</a>	41
<a href="#">2014 级编译原理试题(Done)</a>	44
<a href="#">2015 级编译原理试题(Done)</a>	50
<a href="#">2016 级编译原理试题(Done)</a>	56
<a href="#">2013 级唐敖庆班编译原理试题(Done)</a>	69
<a href="#">2014 级唐敖庆班编译原理试题</a>	75
<a href="#">2015 级唐敖庆班编译原理试题(Done)</a>	78
<a href="#">2016 级唐敖庆班编译原理试题(Done)</a>	81
<a href="#">2017 级唐敖庆班编译原理试题回忆版本</a>	87

Done 标记试题附带部分或者完整的答案,由于时间紧迫,部分试题暂时没有给出参考答案。

## 编译原理知识点总结

一些知识点总结:

### No1. 四元式的翻译

1. 函数调用  $f(X+1, Y)$ ,  $X$  是一般整形变量,  $Y$  是变参整型变量

(+,  $X$ , 1,  $t_1$ )

(VALACT,  $t_1$ , offset, 1)

(VARACT,  $t_1$ , offset + 1, 1)

(CALL,  $f$ , true,  $t_2$ )

2. 函数声明四元式

(ENTRY, LABEL, SIZE, LEVEL)

(ENDPROC, -, -, -)

入口, 标签, AR 大小, 层数

Size 的大小需要回填才知道

3. case 语句四元式?

暂时不知道写

4. for( $E_1$ ; $E_2$ ; $E_3$ ) 语句四元式

$E_1$  中间代码

(FOR, -, -, -)

$E_2$  中间代码

(DO,  $E_2$ .FROM, -, -)

循环的中间代码

$E_3$  中间代码

(ENDFOR, -, -, -)

5. Repeat 语句四元式

(Repeat, -, -, -)

(Until,  $E$ .From, -, -)

(EndRepeat, -, -, -)

### No2. 自底向上的四种方法 LR(0)、SLR(1)、LR(1)、LALR(1)

LR(0) 分析能力低, 局限性大; SLR(1) 容易实现, 但是无法表示一些文法; LR(1) 分析能力最强, 但是实现代价太高; LALR(1) 状态数会比 LR(1) 校一些。

### No3. 归下降程序编写的步骤

1. 构造等价的文法, 消除左递归, 提取公共前缀
2. 求非终极符的 first 和 follow 集
3. 再求文法产生式的 predict 集
4. 再写程序

### No4. 如何把 NFA 化成最小 DFA?

1. 画表格
2. 合并等效状态
3. 画自动机

#### No5. 如何求 first, follow, predict 三个集合?

1. First 集和 follow 集是针对非终极符来说的, predict 集是针对文法产生式来说的。  
对于产生式:  $A \rightarrow BCdE$  来说, First(A) 好求, 直接看这个非终极集有啥产生式, 它就等于产生式能推出的第一个终极符的符号, 如上, 如果 B 的 first 集不空, 则  $\text{First}(A) = \text{First}(B)$ ; 当 B 的 first 集合可以是空, C 的 first 集不会空, 又要加上 First(C); 当 C 的 first 集合可以空的话, 那么  $\text{first}(A) = \{\text{first}(B) - \epsilon\} \cup \{\text{first}(C) - \epsilon\} \cup \{d\}$ ;

2. 求 follow 集合, 建议看右部, 首先求文法开始符的 follow 集, 这个一定含  $\{\#\}$ ; 对于产生式  $S \rightarrow aAb$  来说,  $\text{follow}(A)$  肯定含  $\{b\}$ ; 对于  $X \rightarrow cA$  来说,  $\text{follow}(A)$  含有  $\text{follow}(X)$  的元素; 对于  $Y \rightarrow AC$  来说, 要看 first(C) 可不可以是空, 不可以是空的话  $\text{follow}(A)$  含有 first(C) 的元素; 可以是空的话就是  $\{\text{first}(C) - \epsilon\} \cup \text{follow}(Y)$ ;

3. Predict 集合, 针对产生式来说的。这里分几种情况讨论以下, 对于  $X \rightarrow aA \dots$  来说, 这条产生式的 predict 集合就是  $\{a\}$ , 而对于  $X \rightarrow ABc$  来说, 这条产生式的集合要看 A 和 B 的 first 集合是不是为空, 都是空的话, 那么  $\text{predict}(X \rightarrow ABc) = \{\text{first}(A) - \epsilon\} \cup \{\text{first}(B) - \epsilon\} \cup \{c\}$ ; 如果 A 的 first 不是空, 那么  $\text{predict}(X \rightarrow ABc) = \text{first}(A)$ 。对于  $X \rightarrow \epsilon$  来说,  $\text{predict}(X \rightarrow \epsilon) = \text{follow}(X)$ 。于是就可以知道  $\text{predict}(X \rightarrow A) = \text{first}(A)$  (当 A 的 first 没有  $\epsilon$ ),  $\text{predict}(X \rightarrow A) = \{\text{first}(A) - \epsilon\} \cup \text{follow}(X)$  (A 的 first 集合可以是空)。

容易看出, 求 predict 集, 就要知道 first 集和 follow 集, 于是一般这三个集合总会一起出现, 除非特别熟练可以不画 first 和 follow 集合

#### No6. 结构体变量声明类型声明区分

```
struct{
    int a, b, c;
}s1;                                //结构体类型声明, 还有结构体变量声明 s1

struct Sample1 {
    int a, b, c;
}s1;                                //结构体类型声明
struct Simple1 t1, *t2;             //声明了两个变量

typedef struct {
    int a, b, c;
} Simple2;                          //结构体类型声明
Simple2 t3, *t4;                    //声明了两个变量
```

整个文档都由作者一个人整理, 里面含有敏感题, 请勿外传, 希望资料不要出现在奔腾打印店里。此外, 由于作者是学生, 无法保证答案的正确性, 请多多见谅与包涵。

# 试题样例

1. 已知文法  $G_2$ , 请写出它的递归下降分析程序。  $G_2$ :

$A \rightarrow [B$

$B \rightarrow X] \mid BA$

$X \rightarrow Xa \mid Xb \mid a \mid b$

我的回答: 与  $G_2$  等价的文法  $G_2'$  :

$A \rightarrow [B$

$B \rightarrow X]B'$

$B' \rightarrow AB' \mid \varepsilon$

$X \rightarrow aX'$

$X \rightarrow bX'$

$X' \rightarrow YX' \mid \varepsilon$

$Y \rightarrow a \mid b$

非终极符	First 集	Follow 集
A	{[}	{#, []}
B	{a, b}	{#, []}
B'	{ $\varepsilon$ , []}	{#, []}
X	{a, b}	{[]}
X'	{a, b, $\varepsilon$ }	{[]}
Y	{a, b}	{a, b, []}

注意到第三条产生式和第四条产生式的 predict 集合交集不为空, 于是它不属于 LL(1) 文法。

产生式	Predict 集合
$A \rightarrow [B$	{[}
$B \rightarrow X]B'$	{a, b}
$B' \rightarrow AB'$	{[}
$B' \rightarrow \varepsilon$	{[, #}
$X \rightarrow aX'$	{a}
$X \rightarrow bX'$	{b}
$X' \rightarrow YX'$	{a, b}
$X' \rightarrow \varepsilon$	{[]}
$Y \rightarrow a$	{a}
$Y \rightarrow b$	{b}

## 2003 年编译原理试题

- 1、给出字母表 $\Sigma=\{a, b\}$ 上能识别只有奇数个 a 和奇数 b 的所有符号串的最小 DFA。
- 2、构造一个 DFA，使它恰好能够接受能被 5 整除的二进制数。
- 3、一个句子是否一定有最左（或最右）推导？一个句型是否一定有最左（或最右）推导？如果没有请举例说明原因。
4. 使用驻留法中的跳转方法给出下面类 C 程序（嵌套式）处理结束时的全局符号表。（采用顺序表结构） 表项为层数和偏移

```

void P () {
    int i, j, k;
    void P1 () {
        real m; int n;
        void P2 () {
            bool a, b;
            P2 的过程体
        }
        P1 的过程体
    }
    P 的过程体
}

void Q () {
    real a, b;
    void Q1 () {
        int i, j;
        Q1 的过程体
    }
    Q 的过程体
}
    
```

- 5、分别给出如下程序段优化前和优化后的中间代码，优化只作常量表达式局部优化、公共表达式局部优化和循环优化。其中数组 A 和 B 的定义为：  
array[1..100] of integer; 并假设 integer 类型占 1 个存储单元的空间。

```

a:=0;
j:=(a+1)*5;
while j<100 do
begin
    a:=a+A[j]*B[j];
    if a<10 then k:=0;
    else k:=(x*y+z)/(x*y-z)+(y-z)/(x*y+z);
    j:=j+1;
end
    
```

## 2003 年编译原理试题解答

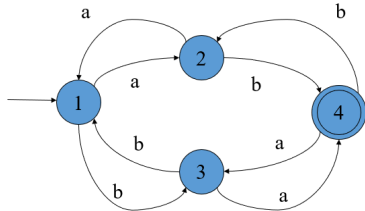
1. 给出字母表识别奇数个 a 和奇数个 b 的所有符号串的最小 DFA  
四个状态：

状态 1：偶数个 a，偶数个 b

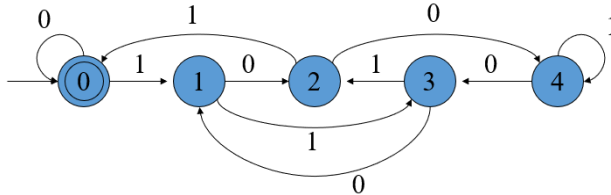
状态 2：奇数个 a，偶数个 b

状态 3：偶数个 a，奇数个 b

状态 4：奇数个 a，奇数个 b



2. 构造一个 DFA，使它恰好能够接受能被 5 整除的二进制数  
除 5 的余数有 0, 1, 2, 3, 4，五个状态



3. 一个句子一定有最左（或最右）推导，但一个句型不一定。

如：有文法 G:

$S \rightarrow AB$

$A \rightarrow a$

$B \rightarrow b$

句型  $S \rightarrow Ab$  不是文法的最左推导

4. 使用驻留法中的跳转方法给出下面类 C 程序（嵌套式）处理结束时的全局符号表。（采用顺序表结构） 表项为层数和偏移。

P	routkind	NULL	actual	0	NULL
i	varkind	intRtr	dir	1	0
j	varkind	intRtr	dir	1	1
k	varkind	intRtr	dir	1	2
P1	routkind	NULL	actual	1	NULL
m	varkind	realRtr	dir	2	0
n	varkind	intRtr	dir	2	2
P2	routkind	NULL	actual	2	NULL
a	varkind	boolRtr	dir	3	0
b	varkind	boolRtr	dir	3	1
#					
#					
#					
Q	routkind	NULL	actual	0	NULL
a	varkind	realRtr	dir	1	0
b	varkind	realRtr	dir	1	2
Q1	routkind	NULL	actual	1	NULL
i	varkind	intRtr	dir	2	0
j	varkind	intRtr	dir	2	1
#					
#					

5、分别给出如下程序段优化前和优化后的中间代码

注意：

除法不外提，赋值不外提

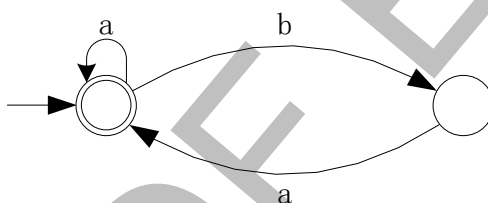
数组相关涉及到优化，不容易被注意到

优化前	优化后
(=, 0, -, a)	(=, 0, -, a)
(+, a, 1, t1)	(=, 5, -, j)
(*, t1, 5, t2)	(*, x, y, t13)
(=, t2, -, j)	(+, t13, z, t14)
(WHILE, -, -, -)	(-, t13, z, t16)
(<, j 100, t3)	(/, t14, t16, t17)
(DO, t3, -, -)	(-, y, z, t18)
(-, j, 1, t4)	(+, t13, z, t20)
(*, t4, 1, t5)	(WHILE, -, -, -)
([], A, t5, t6)	(<, j 100, t3)
(-, j, 1, t7)	(DO, t3, -, -)
(*, t7, 1, t8)	(-, j, 1, t4)
([], B, t8, t9)	(*, t4, 1, t5)
(*, t6, t9, t10)	([], A, t5, t6)
(+, a, t10, t11)	([], B, t5, t9)
(=, t11, -, a)	(*, t6, t9, t10)
(<, a, 10, t12)	(+, a, t10, t11)
(THEN, t12, -, -)	(=, t11, -, a)
(=, 0, -, k)	(<, a, 10, t12)
(ELSE, -, -, -)	(THEN, t12, -, -)
(*, x, y, t13)	(=, 0, -, k)
(+, t13, z, t14)	(ELSE, -, -, -)
(*, x, y, t15)	(/, t18, t20, t21)
(-, t15, z, t16)	(+, t14, t21, t22)
(/, t14, t16, t17)	(=, k, -, t22)
(-, y, z, t18)	(ENDIF, -, -, -)
(*, x, y, t19)	(+, j, 1, t23)
(+, t19, z, t20)	(=, t23, -, j)
(/, t18, t20, t21)	(ENDWHILE, -, -, -)
(+, t14, t21, t22)	
(=, k, -, t22)	
(ENDIF, -, -, -)	
(+, j, 1, t23)	
(=, t23, -, j)	
(ENDWHILE, -, -, -)	

## 2004 级编译原理试题

一、 填空题（每空 2 分，共 40 分）

1. 高级语言的实现方式主要有：\_\_\_\_\_、\_\_\_\_\_和转换方式。
- 2 识别语言  $L(G)=\{a^n \mid n \geq 0\}$  的正则表达式为\_\_\_\_\_。
3. 已知语言  $L=\{a^n b b^n \mid n \geq 1\}$ ，给出能产生该语言的一个上下文无关文法\_\_\_\_\_。
4. 给定文法  $G[S]$   
 $S \rightarrow S \text{ and } S \mid S \text{ or } S \mid \text{not } S \mid p \mid q \mid (S)$   
该文法\_\_\_\_\_（是或不是）二义性文法。
5. 给定如下 DFA，



请写出与其等价的正则表达式\_\_\_\_\_。

6. 给定文法  $G[A]$   
 $A \rightarrow aC$   
 $B \rightarrow dB'$   
 $B' \rightarrow bB' \mid \lambda$   
 $C \rightarrow Abe \mid \lambda$   
则  $\text{Predict}(B' \rightarrow \lambda) =$  \_\_\_\_\_。

7. 设有文法  $G[S]$   
 $S \rightarrow a \mid b \mid (A)$   
 $A \rightarrow SdA \mid S$   
则句型  $(SdSdS)$  的句柄为\_\_\_\_\_。

8. 自顶向下语法分析的条件是对文法中任一非终极符  $A$ ，其任意两个产生式  $A \rightarrow \alpha$  和  $A \rightarrow \beta$ ，都要满足条件\_\_\_\_\_，满足之一条件的文法称为  $LL(1)$  文法。

9. 通过合并  $LR(1)$  文法中的同心状态得到的  $LALR(1)$  文法可能会产生\_\_\_\_\_冲突；一定不会产生\_\_\_\_\_冲突。



10. 假设有如下文法

$S \rightarrow AaAb \mid BbBa$

$A \rightarrow \lambda$

$B \rightarrow \lambda$

该文法\_\_\_\_\_ (是或不是) LL(1) 文法; \_\_\_\_\_ (是或不是) SLR(1) 文法。

11. 假设有下面程序段, 且参数 a 的层数和偏移分别为 L 和 Offset, 假设整数类型、实数类型、地址类型只占一个存储单元, 则参数 b 和 x 的偏移分别为\_\_\_\_\_、\_\_\_\_\_。

```
var at:ARRAY[1..10] OF real;
```

```
function f(a:at; VAR b:at; Var x:real):integer;
```

12. 一个过程 S 在动态链中可有多个过程活动记录 AR, 但其中只有最新 AR(S) 是可访问的, 称此 AR(S) 为 S 的\_\_\_\_\_, 并记作 LiveAR(S)。

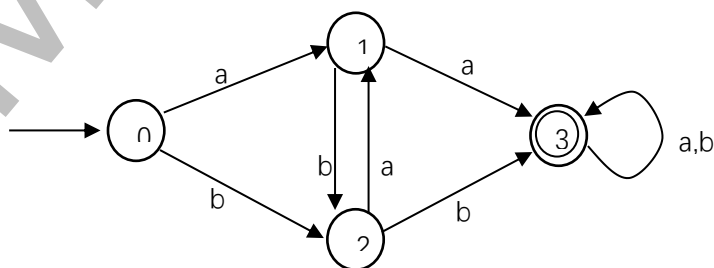
13. 假设有调用链 (M, G, H, R, S), 其中 M, G, H, R, S 的层数分别为 0、1、2、3、3, 并且假设 M, G, H, R, S 对应的活跃活动记录首地址分别为 ar0, ar1, ar2, ar3, ar4, 则过程活动记录 AR(R) 的局部 Display 表 AR(R).Display=\_\_\_\_\_ ; 过程活动记录 AR(S) 的局部 Display 表 AR(S).Display=\_\_\_\_\_。

14. 无符号常数的识别和拼数通常都在编译过程的\_\_\_\_\_阶段完成。

15. 目标代码生成中, 翻译四元式 (THEN, t, \_, \_) 时, 所生成的指令是不完整的, 需要回填; 现假设与该四元式对应的四元式 (ELSE, \_, \_, \_) 所生成的目标指令为 “JMP \_\_\_\_\_”, 且其所在地址为 A, 并假设每条目标指令占用 1 个存储单元, 则此时可以回填 THEN 四元式的目标指令, 请写出回填以后的 THEN 四元式的目标指令\_\_\_\_\_。

二、 简答题 (每题 5 分, 共 25 分)

1. 将如下确定有限自动机化简 (最小化)



2. 处于 /\* 和 \*/ 之间的串构成注释, 注释中间没有 /\*。画出接受这种注释的 DFA 的状态转换图。

2. 什么是过程活动记录? 它主要由哪些内容构成?

4. 已知文法  $G[S]$ :

$S \rightarrow eT \mid RT$

$T \rightarrow DR \mid \varepsilon$

$R \rightarrow dR \mid \varepsilon$

$D \rightarrow a \mid bd$

写出该文法的  $LL(1)$  分析表。

5. 已知文法  $G[S]$

$S \rightarrow aS \mid bS \mid a$

请构造该文法的  $LR(0)$  归约规范活前缀状态机。

三、 综合题 (共 35 分)

1. 已知文法  $G_2$ , 请写出它的递归下降分析程序.

$G_2: A \rightarrow [B$

$B \rightarrow X]BA$

$X \rightarrow Xa \mid Xb \mid a \mid b$

2. 已知文法  $G[S]$

$S \rightarrow A$

$A \rightarrow BA \mid \lambda$

$B \rightarrow aB \mid b$

1) 构造该文法的  $LR(1)$  归约规范活前缀状态机。

2) 构造相应的  $LR(1)$  分析表。

3. 已知如下程序段

`a:=1;`

`while a<=10 do`

`begin`

`if a<>b then`

`A[a][b]:=A[a][b]+2;`

`a:=a+1;`

`end;`

1. 写出生成的四元式中间代码序列;

2. 将中间代码序列划分成基本块, 画出程序流图

3. 写出优化后的中间代码

注: 数组 A: array[1..10] of array[1..10] of integer; 整型类型占 1 个存储单元。

## 2007 级编译原理试题 1

一、选择题 (10 小题, 共 20 分, 每题 2 分)

1. 编译程序必须完成的工作有\_\_\_\_\_。

- ① 词法分析    ② 语法分析    ③ 语义分析  
④ 中间代码生成 ⑤ 中间代码优化    ⑥ 目标代码生成  
A. ①②③④⑤⑥ B. ①②③④ C. ①②③⑥ D. ①②③④⑤

2. 下列关于编译和解释的说法, 正确的是\_\_\_\_\_。

- ① 解释方式和编译方式的区别在于解释程序对源程序并没有进行真正翻译  
② 编译方式与解释方式的根本区别在于是否生成目标代码  
③ 解释程序和编译程序都是语言处理程序  
④ 与编译系统相比, 解释系统比较简单, 可移植性好, 执行速度慢  
⑤ 编译程序是将高级语言程序翻译成汇编语言程序或机器语言程序  
⑥ 解释程序是将汇编语言程序翻译成机器语言程序

- A. ①②③④⑤⑥ B. ①②③④⑤ C. ②③④⑤⑥ D. ②③④⑤

3. \_\_\_\_\_ 这样一些语言, 它们能被确定有限自动机识别, 但不能用正则表达式表示。

- A. 存在 B. 不存在 C. 无法判定是否存在

4. 已知文法  $G[S]$ :

$S \rightarrow SaS \mid SbS \mid ScS \mid dSe \mid f$ ,

下列句型中, \_\_\_\_\_ 是规范句型。

- A.  $fafbS$  B.  $faSbS$  C.  $SaSbf$  D.  $faS$

5. 下列文法中不是二义性文法的是\_\_\_\_\_。

- A.  $\{S \rightarrow +SS \mid -SS \mid a\}$  B.  $\{S \rightarrow S(S)S \mid \varepsilon\}$   
C.  $\{S \rightarrow aSbS \mid bSaS \mid \varepsilon\}$  D.  $\{S \rightarrow a \mid S + S \mid SS \mid S^* \mid (S)\}$

6. 后缀表达式  $ab+c*de*+$  的中缀形式是\_\_\_\_\_。

- A.  $a*(b+c)+d*e$  B.  $(a+b)*c+d*e$  C.  $a+b*c*d+e$  D.  $a+b*c+d*e$

7. 合并无冲突的 LR (1) 状态机得到的 LALR (1) 状态机中一定不会出现\_\_\_\_\_冲突。

- A. 移入/移入冲突 B. 移入/归约冲突 C. 归约/归约冲突

8. 编译程序使用\_\_\_\_\_区别标识符的作用域。

- A. 声明标识符的过程或函数名 B. 声明标识符的过程或函数的静态层数  
C. 声明标识符的过程或函数的动态层数 D. 标识符的行号

9. 适合采用静态存储分配策略的程序设计语言的限制有\_\_\_\_\_。

- ① 数据实体所需空间在编译时能确定 ② 过程调用不允许递归  
③ 不能动态建立数据实体 ④ 运行时每个数据对象只能有一个实例  
⑤ 数组的上下界是常量

- A. ①②③④⑤ B. ①②③④ C. ①②③ D. ①②

10. 目标代码生成器的输出可以是\_\_\_\_\_。

- ① 绝对机器代码 ② 可重定位机器代码 ③ 汇编代码  
④ 虚拟机代码 ⑤ 抽象语法树 ⑥ 三地址代码  
A. ①②③④⑤⑥ B. ①②③④ ⑤ C. ①②③④ D. ①②③

## 二、简答题 (5 小题, 共 30 分, 每个 6 分)

1. 对于文法  $G[S]$ :

$S \rightarrow aAbB$

$A \rightarrow c \mid Ac$

$B \rightarrow d \mid dB$

请画出句型  $aAcbbd$  的语法树, 并求出该句型的所有短语、简单短语和句柄。

2. 写出类型 `Grade07` 的内部表示, 其中每个 `int`、`char` 类型数据各占 1 个内存单元。

```
typedef struct {char name[20]; int num;}student;
```

```
typedef student Grade07[400];
```

3. 请给出下面文法  $G[A]$  的 LL(1) 分析表。

$A \rightarrow aABe [1] \mid Bc [2]$

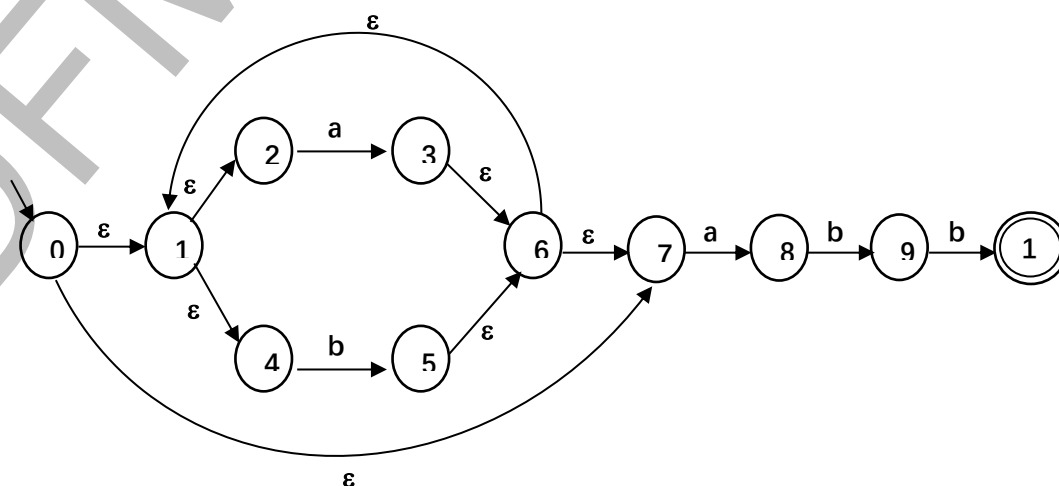
$B \rightarrow dB [3] \mid \varepsilon [4]$

4. 过程活动记录 (AR) 一般应包含哪些信息?

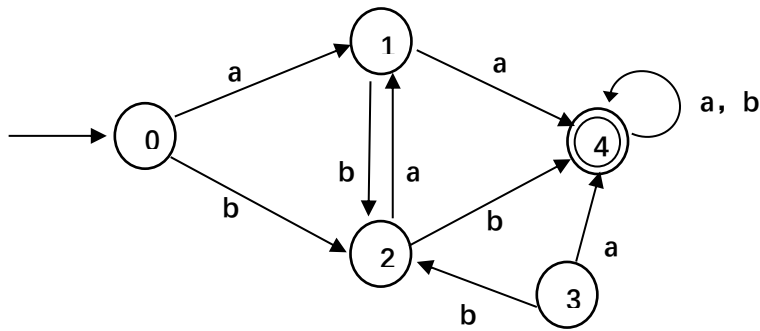
5. 假定当前函数的层数为  $L$ , 偏移量是  $off$ , 每个函数的局部数据区的起始偏移为  $InitOff$ , 每个 `int` 类型数据占 1 个单元。请给出 **A** 至 **F** 位置的层数和偏移量信息。

```
(L, off)
int time = 100; A
int fac(Bint x) { C
    if (x < 0) return -1;
    if ((x == 0) || (x == 1)) return 1;
    else return x*fac(x-1);
} D
void main() {
    int i = 1; E
    while (i <= time) {
        fac(i);
        i++;
    }
} F
```

三、(1) 将下图所示的 NFA  $M$  确定化 (10 分)



(2) 将下图所示的 DFA M 最小化 (5 分)



四、(10 分) 已知文法  $G[D]$  如下:

$D \rightarrow TL$  [1]

$T \rightarrow \text{int}$  [2] |  $\text{real}$  [3]

$L \rightarrow \text{id } R$  [4]

$R \rightarrow , \text{id } R$  [5] |  $\epsilon$  [6]

求出每个产生式的 Predict 集, 并写出该文法的递归下降语法分析程序。

五、(10 分) 已知文法  $G[S]$  如下:

$S \rightarrow eAd$  |  $fBd$  |  $eBa$  |  $fAa$

$A \rightarrow h$

$B \rightarrow h$

判断该文法是否是 LALR(1) 文法, 并说明理由。

六、(15 分) 给出如下程序段的四元式序列:

```

int a[10]={0,1,2,3,4,5,6,7,8,9};
int add(int x[10],int m){
    int s = 0;
    if (m==0) return -1;
    while (m>0)
    {
        s = s + x[10-m];
        m = m -1;
    }
    return s;
}
int main() {
    int sum =0;
    sum = add(a,n);
    if (sum>0) return 0; else return -1;
}
    
```

## 2007 级编译原理试题 2

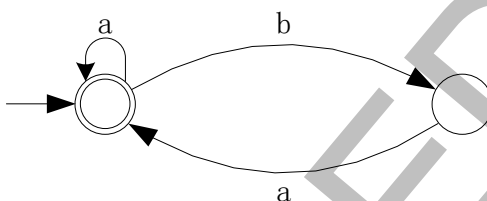
1 设有文法  $G[S]$  如下:

$$S \rightarrow a \mid (T)$$
$$T \rightarrow T, S \mid S$$

请给出句子  $(a, (a, a))$  的所有短语、简单短语和句柄。

2. 请给出描述语言  $L = \{a^n b^i c^j \mid n \geq 1, i \geq 1\}$  的文法。

3. 给定如下 DFA



请写出与其等价的正则表达式。

4. 说明如下文法是否是 LL(1) 文法, 若不是, 将其转换为 LL(1) 文法并给出转换后的文法的 LL(1) 分析表。

$$S \rightarrow (L) \mid aS \mid a$$
$$L \rightarrow LbS \mid S$$

5. 已知文法  $G[S]$

$$S \rightarrow A \quad (1)$$
$$A \rightarrow BA \quad (2)$$
$$\mid \lambda \quad (3)$$
$$B \rightarrow aB \quad (4)$$
$$\mid b \quad (5)$$

1) 构造该文法的 LR(1) 归约规范活前缀状态机。

2) 构造相应的 LR(1) 分析表。

## 2007 级编译原理 2 试题解答

1.

短语:  $(a, (a, a))$ 、 $a, (a, a)$ 、 $a$ 、 $(a, a)$

简单短语:  $a$

句柄:  $a$

注意: 可能要区分第几个  $a$

2.

$S \rightarrow YT$

$Y \rightarrow aYb \mid ab$

$T \rightarrow cT \mid c$

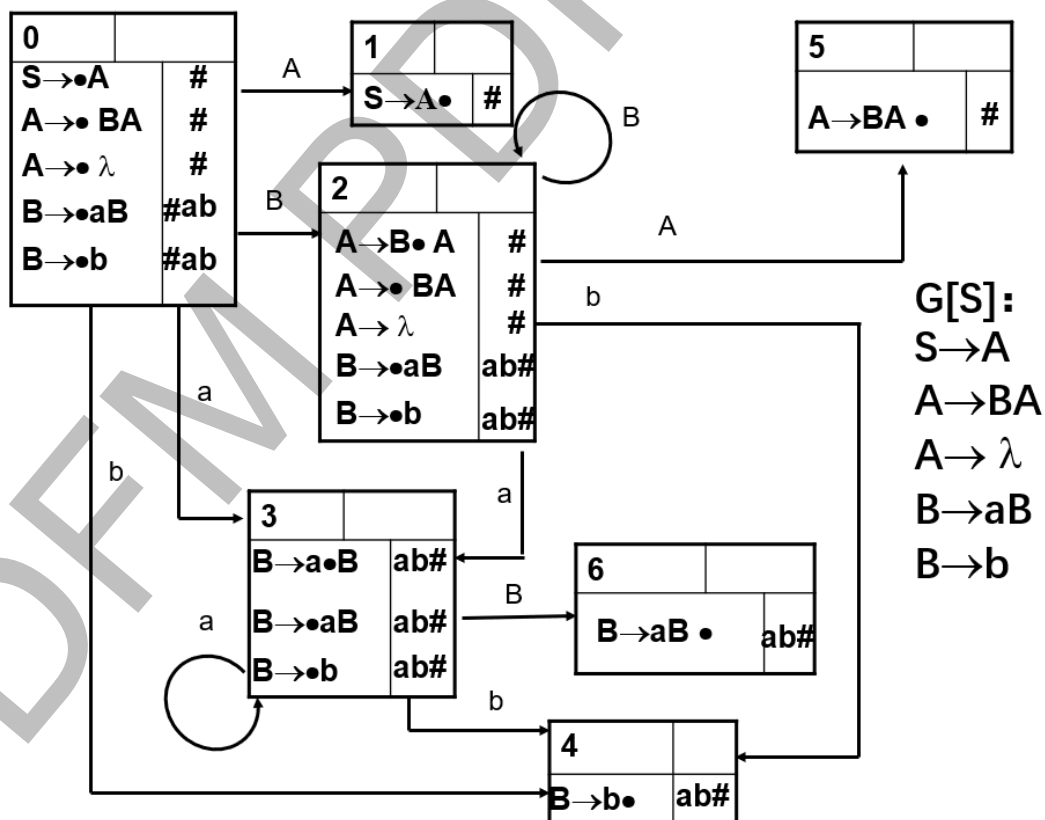
3.  $(a^* \mid (ba))^*$

4. 不是, 首先应该先提取公共前缀, 再消除左递归

	Predict 集
$S \rightarrow (L)$	(
$S \rightarrow aS'$	a
$S' \rightarrow S$	(, a
$S' \rightarrow \varepsilon$	#, b, )
$L \rightarrow SL'$	(, a
$L' \rightarrow bSL'$	b
$L' \rightarrow \varepsilon$	)

LL(1)	(	)	#	a	b
S	1			2	
S'	3	4	4	3	4
L	5			5	
L'		7			6

5.



action				goto		
	a	b	#		B	A
0	S3	S4	R2	0	2	1
1			Acc	1		
2	S3	S4	R2	2	2	5
3	S3	S4		3	6	
4	R4	R4	R4	4		
5			R1	5		
6	R3	R3	R3	6		



## 2008 级编译原理试题

考试时间：2011 年 1 月 13 日

班级 \_\_\_\_\_ 学号 \_\_\_\_\_ 姓名 \_\_\_\_\_

- ✧ 请将答案写在答题纸上，写明题号，不必抄题，字迹工整、清晰；
- ✧ 请在答题纸和试题纸上都写上你的班级，学号和姓名，交卷时请将试题纸、答题纸和草纸一并交上来。

### 一、单项选择题 (10 小题，共 20 分，每题 2 分)

1. “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。”，这种说法正确与否？\_\_\_\_\_。  
A. 正确      B. 不正确      C. 不能确定
2. 下列错误信息中属于语法错误提示信息的是\_\_\_\_\_。  
A. 在数中出现非数字字符  
B. else 没有匹配的 if  
C. 数组下标越界  
D. 使用的函数没有定义
3. 下列文法中，描述语言  $L = \{a^m b^n \mid 0 \leq m \leq 2n\}$  的是\_\_\_\_\_。  
A.  $S \rightarrow AB \quad A \rightarrow aaAb \mid ab \mid \epsilon \quad B \rightarrow Bb \mid \epsilon$   
B.  $S \rightarrow AB \mid aABb \quad A \rightarrow aaAb \quad B \rightarrow Bb$   
C.  $S \rightarrow AASb \mid \epsilon \quad A \rightarrow a \mid aA$   
D.  $S \rightarrow AASb \mid \epsilon \quad A \rightarrow a \mid aA \mid \epsilon$
4. 设有文法  $G[S]: S \rightarrow S1 \mid S0 \mid Sa \mid Sc \mid a \mid b \mid c$ ，下列符号串中\_\_\_\_\_不是该文法的句子。  
A. ab0      B. a0c01      C. aaa      D. bc10
5. 文法  $S \rightarrow AASb \mid \epsilon \quad A \rightarrow a \mid \epsilon$  \_\_\_\_\_二义性文法。  
A. 是      B. 不是
6. 逆波兰式  $xabac+d*e+*+=$  的中缀表示式是\_\_\_\_\_。  
A.  $x=a+b*(a+c)*d+e$       B.  $x=(a+b)*((a+c)*d+e)$   
B.  $x=a+b*((a+c)*d+e)$       D.  $x=(a+b)*(a+c)*d+e$
7. 四元式 (ENDWHILE, -, -, -) 的作用是\_\_\_\_\_。  
A. 标明循环体的结束，转向循环头      B. 标明循环体的结束  
C. 标明循环体      D. 转向循环头

8. 下图所示为某个文法的 LALR(1) 自动机的一个状态, 该状态中是否有冲突? 如果有, 是何种冲突? \_\_\_\_\_。

$S \rightarrow c \bullet a A$	$\{b\}$
$A \rightarrow c \bullet$	$\{b\}$
$B \rightarrow c \bullet$	$\{c\}$

- A. 有冲突, 是移入/移入冲突      B. 有冲突, 是移入/归约冲突  
C. 有冲突, 是归约/归约冲突      D. 无冲突

9. 下列关于文法所描述语言的说法正确的是\_\_\_\_\_。

- A. 文法所描述的语言是文法字母表中符号组成的所有符号串  
B. 文法所描述的语言是文法字母表中所有符号组成的符号串  
C. 文法所描述的语言是文法开始符所能推导出的所有终极符串  
D. 文法所描述的语言是文法开始符推导出的所有符号串

10. 在编译中, 动态存储分配的含义是\_\_\_\_\_。

- A. 在运行阶段对源程序中的量进行分配  
B. 在编译阶段对源程序中的量进行分配  
C. 在编译阶段对源程序中的量进行分配, 在运行时这些量的地址可以根据需要改变  
D. 以上都不正确

二、简答题 (5 小题, 每小题 6 分, 共 30 分)

1. 对于文法  $G[E]$ :

$E \rightarrow E T + \mid T$

$T \rightarrow T F * \mid F$

$F \rightarrow F \wedge \mid a$

请画出句型  $FF \wedge *$  的语法树, 并求出该句型的所有短语、简单短语和句柄。

2. 写出下列标识符的内部表示, 其中每个 `int`、`char` 类型数据各占 1 个内存单元, `float` 类型数据占 2 个内存单元。

```
typedef struct {char name[20]; int age; float score;} student;
student Grade08[100];
```

3. 请给出下面文法  $G[S]$  的各产生式的 Predict 集, 并构造该文法的 LL(1) 分析表。

$S \rightarrow eT \quad [1] \mid RT \quad [2]$

$T \rightarrow DR \quad [3] \mid \varepsilon \quad [4]$

$R \rightarrow dR \quad [5] \mid \varepsilon \quad [6]$

$D \rightarrow a \quad [7] \mid bd \quad [8]$

4. 过程活动记录(AR)中动态链指针、返回地址和返回值的含义是什么, 在过程/函数调用中各起什么作用?

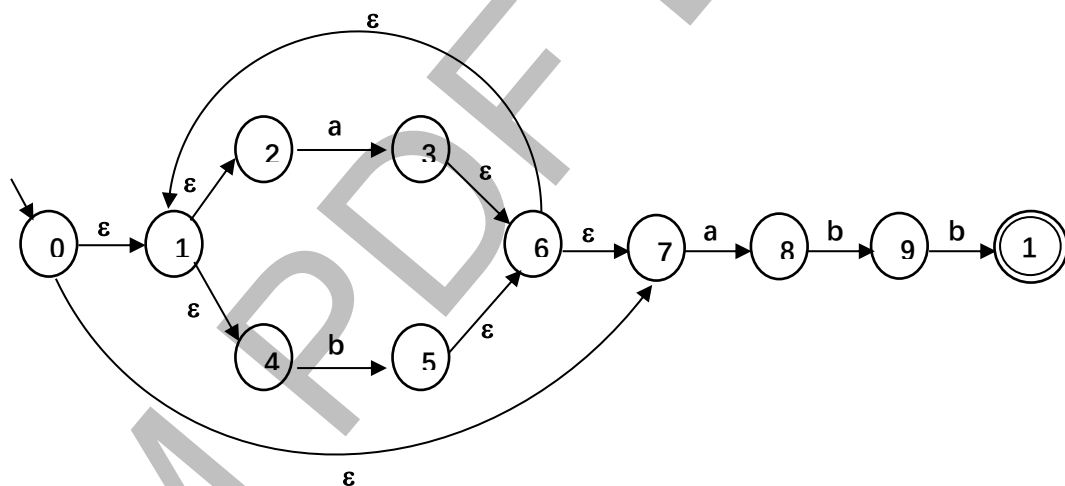
5. 分别给出下面 C 程序扫描到语句“c = a+b+x;”和“d=a+b; ”时相应的符号表内容。

要求：1， 采用局部化符号表方法；

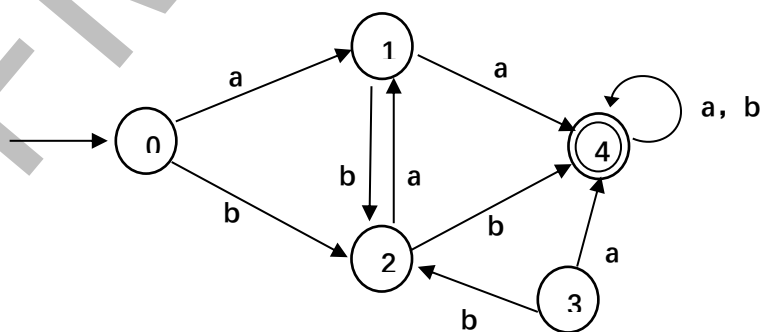
2， 变量标识符要写出层数、偏移量和类型属性， 函数标识符要求写出层数属性；

```
void main ()
{
    int a=0;
    float b=1.0;
    {
        float a=3.0;
        {
            float x=1.3;
            float b=0.3;
            c=a+b+x;
        }
        {
            int b=10;
            d=a+b;
        }
    }
}
```

三、(1) 将下图所示的 NFA M 确定化 (10 分)



a) 将下图所示的 DFA M 最小化 (5 分)



四、(10 分)已知文法  $G[D]$  如下:

$D \rightarrow TL [1]$

$T \rightarrow \text{int} [2] \mid \text{real} [3]$

$L \rightarrow \text{id} R [4]$

$R \rightarrow , \text{id} R [5] \mid \varepsilon [6]$

求出每个产生式的 Predict 集, 并写出该文法的递归下降语法分析程序。

五、(10 分) 已知文法  $G[S]$  如下:

$S \rightarrow eAd \mid fBd \mid eBa \mid fAa$

$A \rightarrow h$

$B \rightarrow h$

判断该文法是否是 LALR(1) 文法, 并说明理由。

六、(15 分) 给出如下程序段的四元式序列:

```
int a[10]={0,1,2,3,4,5,6,7,8,9};
int add(int x[10],int m){
    int s = 0;
    if (m==0) return -1;
    while (m>0)
    {
        s = s + x[10-m];
        m = m -1;
    }
    return s;
}
int main() {
    int sum =0;
    sum = add(a,n);
    if (sum>0) return 0; else return -1;
}
```

# 2010 级编译原理试题

班级 \_\_\_\_\_ 学号 \_\_\_\_\_ 姓名 \_\_\_\_\_

- ✧ 请将答案写在答题纸上，写明题号，不必抄题，字迹工整、清晰；
- ✧ 请在答题纸和试题纸上都写上你的班级，学号和姓名，交卷时请将试题纸、答题纸和草纸一并交上来。

## 一、选择题（2分/题×5题=10分）

1. 文法  $G$  描述的语言  $L(G)$  是指( )。  
A.  $L(G) = \{\alpha \mid S \Rightarrow^+ \alpha, \alpha \in V_T^*\}$     B.  $L(G) = \{\alpha \mid S \Rightarrow^+ \alpha, \alpha \in (V_T \cup V_N)^*\}$   
C.  $L(G) = \{\alpha \mid S \Rightarrow^* \alpha, \alpha \in V_T^+\}$     D.  $L(G) = \{\alpha \mid S \Rightarrow^* \alpha, \alpha \in (V_T \cup V_N)^*\}$
2. 在 C 语言程序中，局部变量 `int p` 存放的位置是\_\_\_\_\_；语句 `p=malloc(sizeof(int)*10)` 申请得到的空间位于\_\_\_\_\_；全局变量 `int globalIndex` 存放的位置是\_\_\_\_\_；局部变量 `static int si` 的存放位置是\_\_\_\_\_。  
①静态区                      ②栈区                      ③堆区                      ④目标代码区  
A. ②④①②                  B. ②④①①                  C. ②③①②                  D. ②③①①
3. 后缀式 `ab+cd+ /` 可用表达式( B )来表示。  
A. `a+b / c+d`                  B. `(a+b) / (c+d)`                  C. `a+b / (c+d)`                  D. `a+b+c / d`
4. 四元式之间的联系是通过( )实现的。  
A. 指针                      B. 临时变量                      C. 符号表                      D. 四元式序号
5. 合并不存在冲突的 LR(1) 项目集的同心集，则( )。  
A. 不会产生新的移进/归约冲突                  B. 不会产生新的归约/归约冲突  
C. 可能会产生新的移进/归约冲突                  D. 可能会产生新的移进/移进冲突

## 二、简答题（5分/题×6题=30分）

1. 构造一个文法  $G$ ，使  $L(G) = \{a^m b^n c^k \mid m=n+k, n \geq 1, m > 1, k \geq 1\}$ 。
2. 已知文法  $G[S]$ :

$S \rightarrow 1A \mid 0B \mid \varepsilon$   
 $A \rightarrow 0S \mid 1AA$   
 $B \rightarrow 1S \mid 0BB$

给出句型 001B 的所有短语、所有简单短语及句柄。

3. 已知文法  $G[Z]$ :

$Z \rightarrow WV$   
 $W \rightarrow aB \mid aW \mid a$   
 $B \rightarrow b \mid bB$   
 $V \rightarrow bV \mid dD$   
 $D \rightarrow d \mid dD$

说明文法  $G[Z]$  为二义性的。

4. 已知文法  $G[S]$ :

$S \rightarrow a \mid b \mid (A)$

$A \rightarrow SdA \mid S$

完成下列表 1 的简单优先关系矩阵, 并判断  $G[S]$  是否为简单优先文法。

表 1.  $G[S]$  简单优先关系矩阵

	a	b	(	)	d	S	A	#
a								
b								
(								
)								
d								
S								
A								
#								

5. 在一个自下而上的分析方法中, 拟采用以下的语法制导翻译模式, 在按一个产生式归约时, 立即执行括号中的动作。

$G[S]$ :

$S \rightarrow bAb$  {print "1" }

$A \rightarrow (B$  {print "2" }

$A \rightarrow a$  {print "3" }

$B \rightarrow Aa$  {print "4" }

若输入序列为  $b(((aa)a)a)b$ , 则求其分析后的输出序列。

6. 写出类型 Grade2010 的内部表示, 其中每个 int、char 类型数据各占 1 个内存单元, float 类型数据占 2 个内存单元。

```
typedef struct {char name[20]; int num; float subj1, subj2;} student;
```

```
typedef student Grade2010[540];
```

三、(12 分) 驻留法实现全局顺序符号表, 给出扫描下述程序后的符号表内容。(每个函数的局部数据区的起始偏移为 initOff, 每个 int 类型数据占 1 个存储单元, 每类名字应包括关键属性。)

```
void main()
{
    int a=1;
    int b=1;
    {
        int b=2;
        {
            int a=3;
            printf("a = %d, b = %d\n", a, b);
        }
        {
            int b = 3;
            printf("a = %d, b = %d\n", a, b);
        }
    }
    printf("a = %d, b = %d\n", a, b);
}
```

四、(12 分)为  $R=(a|b)^*(aa|bb)(a|b)^*$ 构造与其等价的最小 DFA。

五、(12 分)构造下述文法的 LL(1)分析表。

G[S]:

$S \rightarrow S;A \mid A$

$A \rightarrow i := E$

$E \rightarrow E+F \mid F$

$F \rightarrow (E) \mid i$

六、(12 分)已知文法 G[E]:

$E \rightarrow E + T \mid T$

$T \rightarrow TF \mid F$

$F \rightarrow F^* \mid (E) \mid i$

给出该文法LR(1)项目集  $\{[F \rightarrow (\cdot E), \#/+/*/(/i)]\}$  的闭包，以及该项集闭包相对于文法符号 E, T, F 的后继LR(1)项目集闭包。

七、(12 分)已知文法 G[Z]:

1.  $Z \rightarrow CbBA$

2.  $A \rightarrow Aab$

3.  $A \rightarrow ab$

4.  $B \rightarrow c$

5.  $B \rightarrow Db$

6.  $C \rightarrow a$

7.  $D \rightarrow a$

该文法分析表见表 2，试应用 LR (1) 分析技术识别输入符号串 ababab 是否为文法 G[Z]的句子，按表 3 格式写出分析步骤。

8. 表 2 文法 G[Z]的 LR(1)分析表

	action			goto				
	a	b	#	Z	A	B	C	D
0	$S_3$			1			2	
1			Accept					
2		$S_4$						
3		$R_6$						
4	$S_8$					5	6	7
5	$S_{11}$				10			
6	$R_4$							
7		$S_9$						
8	$R_6$	$R_7$						
9	$R_5$							
10	$S_{13}$		$R_1$					
11		$S_{12}$						
12	$R_3$		$R_3$					
13		$S_{14}$						
14	$R_2$		$R_2$					

表 3 符号串 ababab 的分析步骤

步骤	状态栈	符号栈	输入流	action	goto
1	0		ababab#	shift	3
2					
.....	.....	.....	.....	.....	.....



2013-2014 学年 第 2 学期

## 2011 级《编译原理》期末考试试题(A 卷)

考试时间: 2014 年 7 月 1 日

班级 \_\_\_\_\_ 学号 \_\_\_\_\_ 姓名 \_\_\_\_\_

- ◇ 请将答案写在答题纸上, 写明题号, 不必抄题, 字迹工整、清晰;  
◇ 请在答题纸和试题纸上都写上你的班级, 学号和姓名, 交卷时请将试题纸、答题纸和草纸一并交上来。

一、选择题 (2 分/题  $\times$  10 题=20 分), 从下面各题所给的 A、B、C、D 四个选项中, 选择最佳选项填入空白处。

1. 文法  $G: S \rightarrow xSx \mid y$  所识别的语言是\_\_\_\_\_。  
A.  $\{xyx\}$       B.  $(xyx)^*$       C.  $\{x^n y x^n \mid n \geq 0\}$       D.  $x^* y x^*$
2. 词法识别器的输入是\_\_\_\_\_。  
A. 单词符号      B. 源程序      C. 语法单位      D. 目标程序
3. 在词法分析过程中能够发现\_\_\_\_\_错误。  
A. 操作数类型不匹配      B. 标识符重复声明  
C. 程序中出现非法符号      D. 除法溢出
4. 作为语法分析器子程序时, 词法识别器的返回结果是\_\_\_\_\_。  
A. 单词属性值      B. 单词的行号和语义值  
C. 单词的种别编码和语义值      D. 单词的种别编码
5. DFA  $M_1$  和 DFA  $M_2$  等价是指\_\_\_\_\_。  
A.  $M_1$  和  $M_2$  的状态集相等  
B.  $M_1$  和  $M_2$  的状态集和状态转换函数分别相等  
C.  $M_1$  和  $M_2$  所识别的字符串集合相等  
D.  $M_1$  和  $M_2$  的状态转换函数相等
6. 编译程序常常先把源代码转换成中间代码, 然后再将中间代码翻译成目标代码, 这样做的好处是\_\_\_\_\_。  
①利用有限的机器内存并提高机器的执行效率  
②为使编译程序在逻辑结构上更为简洁  
③可以在中间代码一级进行优化工作, 使得目标代码的生成比较容易  
④便于编译程序的移植, 将与机器相关的某些实现细节置于代码生成阶段进行处理  
A. ③④      B. ②③④      C. ①②      D. ①②④

7. 文法  $G$  的句型是指\_\_\_\_\_。

- A.  $\{\alpha | S \Rightarrow^+ \alpha, \alpha \in V_T^*\}$       B.  $\{\alpha | S \Rightarrow^+ \alpha, \alpha \in (V_T \cup V_N)^*\}$   
 C.  $\{\alpha | S \Rightarrow^* \alpha, \alpha \in V_T^*\}$       D.  $\{\alpha | S \Rightarrow^* \alpha, \alpha \in (V_T \cup V_N)^*\}$

8. 后缀式  $xabac-d*e-*$  可用于描述中缀式\_\_\_\_\_。

- A.  $x=a-b*((a-c)*d-e)$       B.  $x=(a-b)*((a-c)*d-e)$   
 C.  $x=a-b*(a-c)*d-e$       D.  $x=(a-b)*(a-c)*d-e$

9. 关于 LR(0)、SLR(1)、LR(1) 及 LALR(1) 四种 LR 类语法分析方法, 下列描述正确的是\_\_\_\_\_。

- ①从分析能力方面看, 有:  $LR(0) < SLR(1) < LALR(1) < LR(1)$   
 ②从分析能力方面看, 有:  $LR(0) < LR(1) < SLR(1) < LALR(1)$   
 ③从活前缀状态机的状态数量方面看, 有:  $LR(0) < SLR(1) < LALR(1) < LR(1)$   
 ④从活前缀状态机的状态数量方面看, 有:  $LR(0) = SLR(1) = LALR(1) < LR(1)$

- A. ①④      B. ①③      C. ②③      D. ②④

10. 在下述语法分析方法中, 属于自底向上的是\_\_\_\_\_。

- ①LL(1)      ②SLR(1)      ③LR(0)      ④LR(1)  
 ⑤LALR(1)      ⑥递归下降子程序方法      ⑦简单优先分析方法
- A. ①②③④      B. ①②③  
 C. ②③④⑤⑦      D. ③④⑤⑦

## 二、简答题 (5 分/题×8 题=40 分)

1. 已知文法  $G[E]$ :

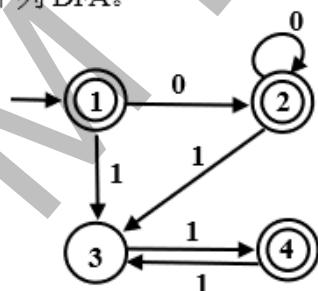
$E \rightarrow E+T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid i$

给出句型  $E+T * F+i$  的所有短语、所有简单短语及句柄。

2. 最小化下列 DFA。



3. 已知文法  $G[Z]$ :

$$\begin{aligned} Z &\rightarrow WV \\ W &\rightarrow aB \mid aW \mid a \\ B &\rightarrow b \mid bB \\ V &\rightarrow bV \mid dD \\ D &\rightarrow d \mid dD \end{aligned}$$

试说明  $G[Z]$  为二义性文法。

4. 已知文法  $G[S]$ :

$$\begin{aligned} S &\rightarrow (SR \\ S &\rightarrow a \\ R &\rightarrow ;SR \\ R &\rightarrow ) \end{aligned}$$

试按下列表 1 所示格式完成  $G[S]$  的简单优先关系矩阵。

表 1.  $G[S]$  简单优先关系矩阵

	S	R	(	a	;	)	#
S							
R							
(							
a							
;							
)							
#							

5. 设有 C 语言类型声明:

```
typedef struct {char name[15]; int no; float computer;} STUDENT;
typedef STUDENT SOFTWARE[300];
```

试写出类型 SOFTWARE 的内部表示, 其中每个 int、char 类型数据各占 1 个内存单元, float 类型数据占 2 个内存单元。

6. 设有 C 语言变量声明:

```
int i,k;
float a[10][5];
```

试写出下标变量  $a[i][k]$  的中间代码, 每个 float 类型数据占 2 个存储单元。

7. 令  $L = \{SaS^r \mid S \in (0|1)^*, S^r \text{ 表示 } S \text{ 的逆, 即: 若 } S=110, \text{ 则 } S^r=011\}$ , 试给出语言 L 的上下文无关文法表示。要求: 文法产生式的个数不超过 3 条。

8. 设当前层数为 L，可用偏移为 off，每个函数第一个形参可用的偏移为 m，每个 int 类型数据和指针类型数据各占 1 个存储单元，每个 float 类型数据占 2 个存储单元。试写出①至⑩各程序点的层数和偏移。

```

① typedef int AT[5][10];
   typedef struct {int number; AT score;} BT;
② AT x,y;
   float u,v;
③ int f (AT a, ④AT * b, ⑤float x)
    { ⑥ BT m,n;
      ⑦ int am,an;
        ..... }⑧
void g ( )
{ ⑨
  int k;
  ..... } ⑩
    
```

三、(12 分)采用驻留法实现全局线性组织的符号表，给出扫描下述程序段后的符号表内容。（注：每个函数第一个形参可用的偏移为 m，每个 int 类型数据占 1 个存储单元，每类名字应包括关键属性。）

```

void main()
{
    int a=1;
    int b=1;
    {
        int b=2;
        {
            int a=3;
            printf("a = %d, b = %d\n",a, b);
        }
        {
            int b = 3;
            printf("a = %d, b = %d\n",a, b);
        }
    }
    {
        int a = 2;
        printf("a = %d, b = %d\n",a, b);
    }
    printf("a = %d, b = %d\n", a, b);
}
    
```

四、(12 分)已知文法 G:

$$E \rightarrow E-T \mid T$$

$$T \rightarrow T*F \mid F$$

$$F \rightarrow @P \mid P$$

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

(1) 构造一个与 G 等价的 LL(1)文法 G';

(2) 对于 G', 构造其 LL(1)分析表。

五、(16 分)已知文法 G[Z]:

1.  $Z \rightarrow A$

2.  $A \rightarrow BA$

3.  $A \rightarrow \epsilon$

4.  $B \rightarrow aB$

5.  $B \rightarrow b$

(1) 求 LR(1)项目集  $\{[Z \rightarrow \cdot A, \#]\}$  的闭包  $IS_0$ , 以及  $GO(IS_0, a)$  和  $GO(IS_0, B)$ ;

(2) 该文法 LR(1)分析表见表 2, 试应用 LR(1)分析技术识别输入符号串 ab, 要求按表 3 格式写出分析过程。

表 2 文法 G[Z]的 LR(1)分析表

	action			goto		
	a	b	#	Z	A	B
0	S <sub>3</sub>	S <sub>4</sub>	R <sub>3</sub>		1	2
1			Accept			
2	S <sub>3</sub>	S <sub>4</sub>	R <sub>3</sub>		5	2
3	S <sub>3</sub>	S <sub>4</sub>				6
4	R <sub>5</sub>	R <sub>5</sub>	R <sub>5</sub>			
5			R <sub>2</sub>			
6	R <sub>4</sub>	R <sub>4</sub>	R <sub>4</sub>			

表 3 符号串 abab 的分析过程

步骤	状态栈	符号栈	输入流	动作	转向
1	0	#	ab#	shift	3
2	03	#a	b#	shift	4
.....	.....	.....	.....		

## 2012 级编译原理试题

### 一、选择题

1. 有文法 G:  $E \rightarrow E * T \mid T$

$T \rightarrow T + I \mid i$

若  $1+2*8+6$  按该文法归约, 则其值为 ( )。

A. 23                  B. 42                  C. 30                  D. 17

2. 若符号表保留至目标代码生成阶段, 则表达式  $x+y$  的中间代码  $(+, x, y, t1)$  和部分目标代码  $(ADD R, x)$  中的  $x$  分别代表什么含义 ( )。

①变量  $x$       ② $x$  在符号表中的位置      ③ $x$  的值      ④ $x$  在内存中的相对地址

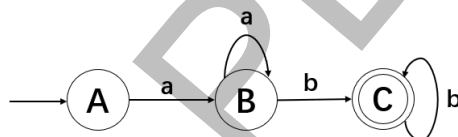
A. ①②                  B. ①③                  C. ②③                  D. ②④

3. 对于 LL(1) 和 LR(1) 分析法输入流中的非终极符 ( ) 符号栈或分析栈。

A. 都移入                          B. 都不移入  
C. 前者移入, 后者不移入      D. 前者不移入, 后者移入

4. 有语言  $L = \{a^n b^n \mid n > 0\}$ , 则下列关于  $L$  的描述正确的是 ( )。

A. 识别  $L$  的正则表达式可以是  $aa^*bb^*$   
B. 识别  $L$  的 DFA 可以是下图所示自动机



C. 识别  $L$  的正则文法可以是

$A \rightarrow aB$

$B \rightarrow aB \mid bC \mid b$

$C \rightarrow bC \mid b$

D. 识别  $L$  的 2 型文法可以是  $S \rightarrow aSb \mid ab$

5. 下列关于过程活动记录的相关描述错误的是 ( )。

A. 运行时过程活动记录的管理工作是由编译器生成的目标代码运行时完成的  
B. 若目标代码确定, 则目标代码生成阶段可以精确计算出每个过程活动记录的大小  
C. 同一个函数被多次调用时所生成的所有过程活动记录的大小一定是相同的  
D. 存在间接递归时, 递归调用的函数的多个过程活动记录可能不止一个是活跃的

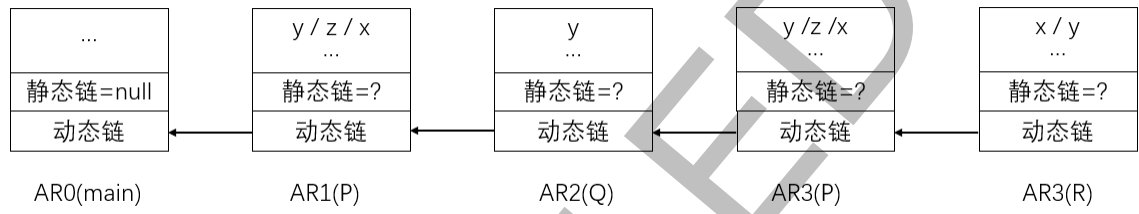
二、简答题

1.构造一个 2 型文法  $G$ ，使得  $L(G)=\{a^{2m-1}b^m \mid m > 0\}$ 。

2.设有如下含有嵌套定义类 C 程序，

```
int P(){
    int y, z, x;
    int Q(){ int y;  Q 的函数体; }
    int R(){ int x;  int y;  x = x + y + z; }
    P 的函数体;
}
void main(){ 主函数体; }
```

若有调用链(main, P, Q, P, R)，则运行时 5 个活动记录（省略部分内容）示意图见下图。



请回答，AR1~AR4 的静态链指针分别指向哪一个活动记录？语句  $x = x + y + z$  使用的  $z$  在哪一个活动记录中。

3.设有语句：

```
do{ S1;
   do{ S2;
      }until(x > 0);
   }until(y > 0);
```

编写其目标代码，S 的目标代码用 S.code 表示。

4.有如下文法的句型  $SmBa(SmA)$ ，请给出该句型的简单短语。

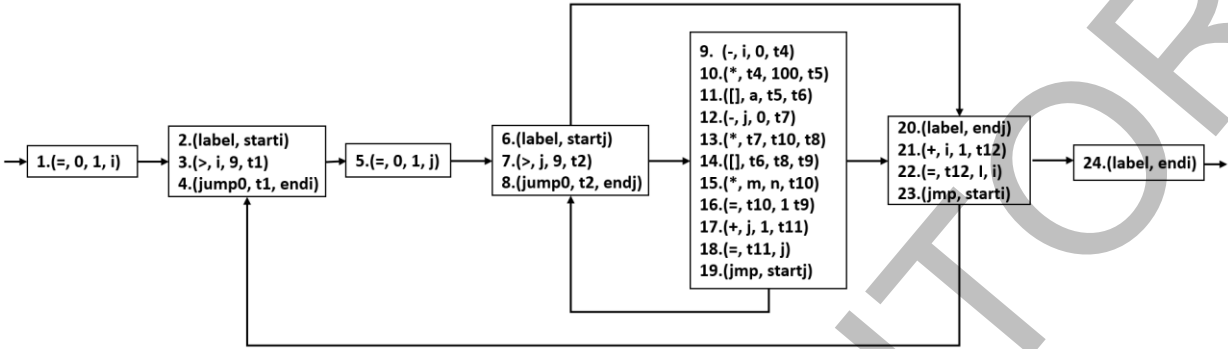
```
S → SmA | A
A → AaB | B
B → (S) | a
```

5.某指令集中，算数运算指令名的命名规则如下（双引号不是指令名的组成部分）：

- a)以 “\_m128\_” 开头
- b)以 “p” 或 “s” 或空结尾
- c)中间是长度大于 0 的任意小写字母构成的串。

请给出识别该指令集中算数运算指令命名的正则表达式。

6.图所示为某两重循环程序段生成的由基本块为结点的四元式程序流图，请写出循环不变式外提优化后的四元式（只需按照顺序写出四元式原有编号即可）。



7. “不含公共前缀和左递归的 2 型文法一定是 LL(1)文法”，这种说法是否正确？若正确请证明之，否则请举例说明。

8.设有文法 G:

$S \rightarrow 1A0B$   
 $A \rightarrow 1A \mid 1$   
 $B \rightarrow 0B \mid 10$

请在文法中插入相应的语义动作（并给出动作代码）完成以下功能：对输入文法 G 的句子，求其含有多少个 1.

9.下面是计算 Fibonacci 数列第 m 项值的 C 函数，请给出语义分析至 return f2 时的全局符号表。

```
int fib(int m){
    int f0, f1, f2;
    f0 = 0; f1 = 1;
    if(m <= 1)    return m;
    else{
        int i;
        for(int i = 2; i <= m; i++){
            f2 = f0 + f1;    f0 = f1;    f1 = f2;
        }
        return f2;
    }
}
```



10.下面是 if 语句的二义性文法：

- $Z \rightarrow S$  (0)  
 $S \rightarrow iSeS$  (1)  
 $S \rightarrow iS$  (2)  
 $S \rightarrow S;S$  (3)  
 $S \rightarrow a$  (4)

假设规定“e 与最近的 i 相匹配，终极符服从左结合”，则可以构造出该文法的无冲突的 SLR(1)分析表，如右图所示。

	action					goto
	i	e	;	a	#	
0	S3			S2		1
1			S4		Acc	
2		R4	R4		R4	
3	S3			S2		5
4	S3			S2		6
5		S7	S4		R2	
6		R3	R3		R3	
7	S3			S2		8
8		R1	S4		R1	

请问用 SLR(1)算法对句子 iiaea#进行分析过程中，

在查表得到 Acc 前共进行了几次移入、几次归约动作？

三、有如下文法：

- $Z \rightarrow aAc$   
 $A \rightarrow bB \mid ba$   
 $B \rightarrow dB \mid e$

请给出识别该文法的所有归约规范活前缀的最小 DFA。

四、已知一个识别语言  $L=\{a^m b^n \mid n > m \geq 0\}$  的 2 型文法如下：

- $Z \rightarrow S$   
 $S \rightarrow aSb \mid B$   
 $B \rightarrow Bb \mid b$

(1) 构造该文法的 LR(1)归约规范活前缀状态机；

(2) 该文法是否为 LR(1)文法？若是请构造 LR(1)分析表，若不是请说明理由。

六、有类 C 语言程序：

```

real F(real k){ return k + k; }
void Q(real x){
    real u, n[5];
    n[3] = 10; u = F(50);
    while(u > n[3])
        u = u - 2;
}
  
```

写出其中间代码。

注：return 语句中间代码为(Return, -, -, var)

## 2012 级编译原理试题解答

### 一、选择题

1~5 BDDDD

### 二、简答题

1.  $S \rightarrow aaSb \mid ab$

2. null AR1 null AR3 AR3

3. 目标代码如下：

[1] S1.code  
[2] S2.code  
[3] LD R, x  
[4] GT R, 0  
[5] JUMP0 R, [2]  
[6] LD R, y  
[7] GT R, 0  
[8] JUMP0 R, [1]

4. SmA、B

5.  $\_m128\_ [a-z]([a-z])^*(p|s|\epsilon)$

6. 优化后的编号如下（从左往右，从上到下的顺序）：

1 15 2 3 4 5 9 10 11 6  
7 8 12 13 14 16 17 18 19 20  
21 22 23 24

7. 错误，反例：  $S \rightarrow AS \mid a$        $A \rightarrow a$

上面的文法没有 公共前缀和左递归，但是却不是 LL(1)文法，所以说判断一个文法是不是 LL(1)文法，就看同一个非终极符 predict 的交集是不是空。

8. 语法制导很重要，是实现编译的核心技术之一，这个题的语法制导很简单，与教材例子类似。

$S \rightarrow \#init\#1\#ADD\#A0B$

$A \rightarrow 1\#ADD\#A \mid 1\#ADD\#$

$B \rightarrow 0B \mid 1\#ADD\#0$

其中，init：初始化 sum=0；ADD：sum++

9. 尽量写出完整的符号表，下面的符号表变量缺少 Value 项，函数缺少 Code 和 Size 项。

Name	Kind	Type	Access/class	Level	ParamList/offset	forward
fib	rouKind	intPtr	dir	0		false
m	varKind	intPtr	dir	1	0	
f0	varKind	intPtr	dir	1	1	
f1	varKind	intPtr	dir	1	2	
f2	varKind	intPtr	dir	1	3	
i	varKind	intPtr	dir	1	4	
#						
#						

10. 5 次移入，4 次归约

三、这个题目考的知识点涉及派生定理：

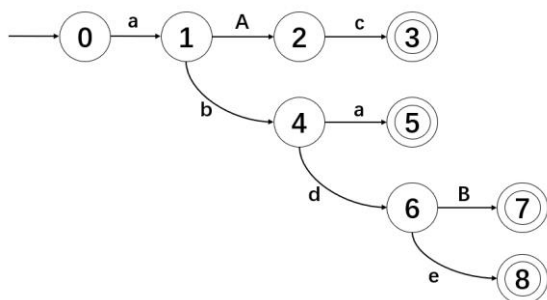
- 1) 开始符产生式的右部是归约活前缀；
- 2) 如果  $\alpha A \beta$  是归约活前缀，且  $A \rightarrow \pi$  是产生式，则  $\alpha \pi$  也是归约活前缀；
- 3) 任何归约活前缀，都可按上述方法被派生。

回归这道题，所有归约规范活前缀如下：

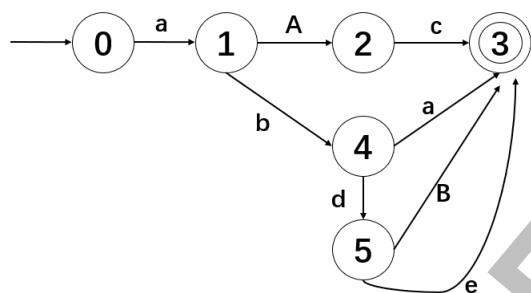
aAc    abB    aba    abdB    abe    abddB    abde    .....

总结：aAc、aba、abd\*B、abd\*e

于是可以画出状态转换图，如下：



正常来说求最小 DFA 是需要画表然后合并状态的，对于这题，很容易看出来状态 3，5，7，8 是可以合并成一个状态的，其他的不能合并。于是，最小 DFA 如下：



点评：这道题考的知识点涉及到了派生定理，可能比较偏，需要引起重视。同时，对于那些想复习完所有知识点的，给个提示：如何求某个文法的规范句型的规范活前缀？这个知识点比较容易忽略。教材 P70 给了一个例题，请钻研弄明白。

四、

**待写**

五、等价文法如下：

$S \rightarrow (L)$  (1)

$S \rightarrow a$  (2)

$L \rightarrow SL'$  (3)

$L' \rightarrow SL'$  (4)

$L' \rightarrow \epsilon$  (5)

LL(1)分析表如下：

	(	)	a	=	#
S	1		2		
L	3		3		
L'		5		4	

六、中间代码如下：

(Entry, F, -, -)

(+, k, k, t1)

(Return, -, -, t1)  
(Endfunc, -, -, -)  
(Entry, Q, -, -)  
(-, 3, 0, t2)  
(\*, t2, 2, t3)  
([], n, t3, t4)  
(=, 10, -, t4)  
(Varact, 50, offset, 2)  
(Call, F, true, t5)  
(=, t5, -, u)  
(While, -, -, -)  
(-, 3, 0, t6)  
(\*, t6, 2, t7)  
([], n, t7, t8)  
(>, u, t8, t9)  
(Do, t9, -, -)  
(-, u, 2, t10)  
(=, t10, -, u)  
(Endwhile, -, -, -)  
(Endfunc, -, -, -)

## 2013 级编译原理试题 1

1 已知文法  $G[S]$ :

$S \rightarrow 1A \mid 0B \mid \varepsilon$

$A \rightarrow 0S \mid 1AA$

$B \rightarrow 1S \mid 0BB$

给出句型 001B 的所有短语、所有简单短语及句柄。 (5.0 分)

2 已知文法  $G[S]$ :

$S \rightarrow a \mid b \mid (A)$

$A \rightarrow SdA \mid S$

完成下列表 1 的简单优先关系矩阵，并判断  $G[S]$  是否为简单优先文法。

表 1.  $G[S]$  简单优先关系矩阵

	a	b	(	)	d	S	A	#	(5.0 分)
a									
b									
(									
)									
d									
S									
A									
#									

3 写出类型 Grade2010 的内部表示，其中每个 int、char 类型数据各占 1 个内存单元, float 类型数据占 2 个内存单元。

```
typedef struct {char name[20]; int num; float subj1, subj2;} student;
typedef student Grade2010[540]; (5.0 分)
```

4、已知文法 G[Z]:

1.  $Z \rightarrow CbBA$
2.  $A \rightarrow Aab$
3.  $A \rightarrow ab$
4.  $B \rightarrow c$
5.  $B \rightarrow Db$
6.  $C \rightarrow a$
7.  $D \rightarrow a$

该文法分析表见表 2，试应用 LR (1) 分析技术识别输入符号串 ababab 是否为文法 G[Z]的句子，按表 3 格式写出分析步骤。

表 2 文法 G[Z]的 LR(1)分析表

	action			goto				
	a	b	#	Z	A	B	C	D
0	S <sub>3</sub>			1			2	
1			Accept					
2		S <sub>4</sub>						
3		R <sub>6</sub>						
4	S <sub>8</sub>					5	6	7
5	S <sub>11</sub>				10			
6	R <sub>4</sub>							
7		S <sub>9</sub>						
8	R <sub>6</sub>	R <sub>7</sub>						
9	R <sub>5</sub>							
10	S <sub>13</sub>		R <sub>1</sub>					
11		S <sub>12</sub>						
12	R <sub>3</sub>		R <sub>3</sub>					
13		S <sub>14</sub>						
14	R <sub>2</sub>		R <sub>2</sub>					

5. 已知文法 G[E]:

- $$\begin{aligned}
 E &\rightarrow E + T \mid T \\
 T &\rightarrow TF \mid F \\
 F &\rightarrow F^* \mid (E) \mid i
 \end{aligned}$$

给出该文法 LR(1)项目集  $\{[F \rightarrow (E), \#/+/*/(i)]\}$  的闭包，以及该项集闭包相对于文法符号 E, T, F 的后继 LR(1)项目集闭包。(15.0 分)

2013 级编译原理试题解答

1. 已知文法  $G[S]$ :

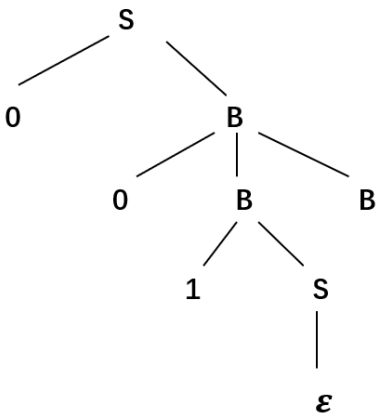
$S \rightarrow 1A \mid 0B \mid \varepsilon$

$A \rightarrow 0S \mid 1AA$

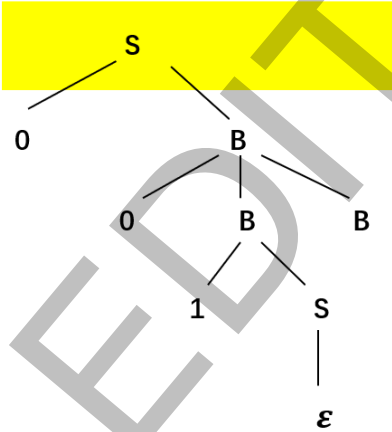
$B \rightarrow 1S \mid 0BB$

给出句型 001B 的所有短语、所有简单短语及句柄

这种类型题，先画语法生成树，然后采取从顶向下观察出短语。



图一



图二

短语: 001B 01B 1  $\varepsilon$

简单短语:  $\varepsilon$

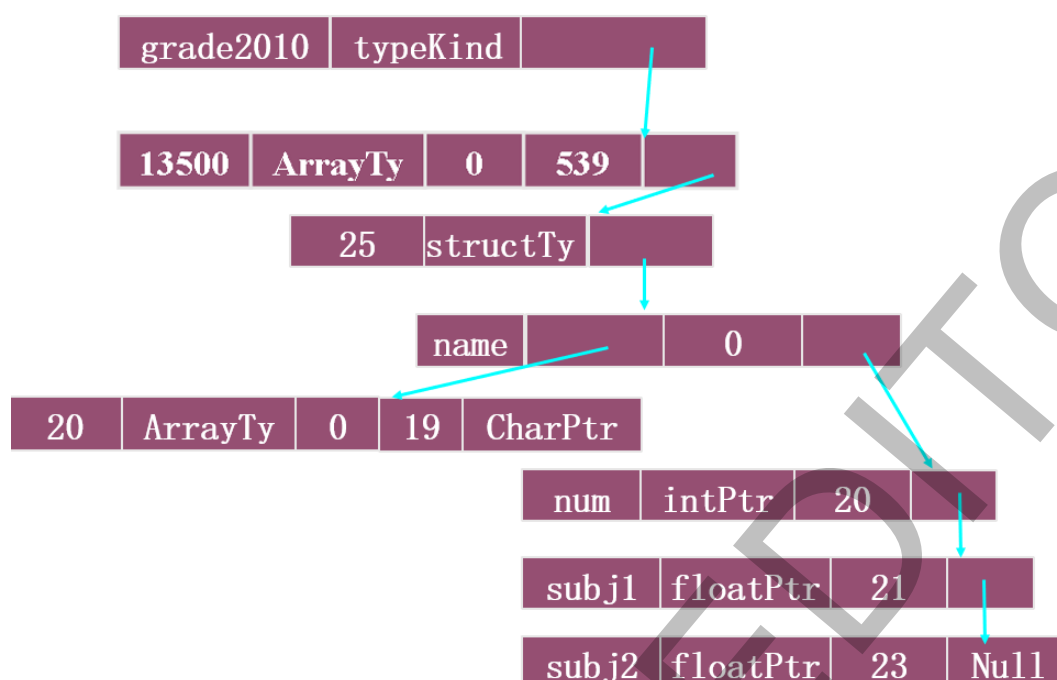
句柄:  $\varepsilon$

2.

	S	A
F	a,b,(	a,b,(
L	a,b,)	a,b,)

	a	b	(	)	d	S	A	#
a				▷	▷			▷
b				▷	▷			▷
(	▷	▷	▷			▷	≅	
)				▷	▷			▷
d	▷	▷	▷			▷	≅	
S				▷	≅			▷
A				≅▷				
#	▷	▷		▷		▷		

3.



4. 2010 年软件学院编译原理一样的题，解答省略。

5. 已知文法  $G[E]$ :

$E \rightarrow E + T \mid T$

$T \rightarrow TF \mid F$

$F \rightarrow F* \mid (E) \mid i$

给出该文法 LR(1) 项目集  $\{[F \rightarrow (\bullet E), \#/+/*/(/i)]\}$  的闭包，以及该项集闭包相对于文法符号  $E, T, F$  的后继 LR(1) 项目集闭包。

主要考的是投影运算、闭包运算

$F \rightarrow (\bullet E)$	$\# + * (i$	$T \rightarrow \bullet F$	$(+) i$
$E \rightarrow \bullet E + T$	$) +$	$F \rightarrow \bullet F*$	$*)+(i$
$E \rightarrow \bullet T$	$) +$	$F \rightarrow \bullet (E)$	$*)+(I$
$T \rightarrow \bullet TF$	$(+) i$	$F \rightarrow \bullet i$	$*)+(i$

输入T	$E \rightarrow T \bullet$	$) +$	$F \rightarrow \bullet F*$	$* (i$
	$T \rightarrow T \bullet F$	$(i$	$F \rightarrow \bullet (E)$	$(i*$
			$F \rightarrow \bullet i$	$(i*$

输入E	$F \rightarrow (E \bullet)$	$\# + * (i$
	$E \rightarrow E \bullet + T$	$) +$

输入F	$T \rightarrow F \bullet$	$(+) i$
	$F \rightarrow F \bullet *$	$*)+(i$



## 2013 级编译原理试题 2

### 一、简答题

1. 典型的编译程序在逻辑功能上由哪几部分组成？

2. 设非确定有限自动机  $M=(S=\{x,y\}, \Sigma=\{a,b\}, f, x, \{y\})$ ，其中  $f$  的定义如下：

$$f(x,a)=\{x,y\}, f(x,b)=\{y\}, f(y,b)=\{x,y\}.$$

请构造与  $M$  等价的最简确定有限自动机。

3. 已知语言  $L=\{aibj \mid j \geq i \geq 1\}$ ，试构造产生该语言的上下文无关文法。

4. 设有文法  $G[S]$  如下：

$$S \rightarrow (L) \mid aS \mid a$$

$$L \rightarrow L, S \mid S$$

请给出句型  $(S, (a))$  的所有短语、简单短语和句柄。

5. 给出下面 C 程序扫描到语句 “ $b=a+b;$ ” 时相应的符号表内容。

要求：1. 采用全局符号表驻留方法；

2. 标识符具有属性信息只需写出名字、层数和偏移。

```
void main
{
    int a = 0;    int b = 1;
    {
        float a = 3.0;
        {
            float x = 1.3;    float b = 0.3;    a = a + b + x;
        }
        {
            int b = 10;
        }
        b = a + b;
    }
}
```

6. 对于下面的 C 语言程序，画出程序运行时第二次递归进入函数  $f$  后的运行时栈的内容是什么？

要求：1. 详细刻画各活动记录数据信息包含的变量及其值；

2. 管理信息的详细内容可以不用给出。

```

int g(int s){
    s = s + 10;
    return s;
}
int f(int n){
    if(n <= 0)
        return 1;
    else
        return n * f(n - 1);
}
void main(){
    int k = 1;
    k = g(20);
    k = f(10);
}

```

二、已知文法  $G[P]$  如下：

- [1]  $P \rightarrow \{d;S\}$
- [2]  $S \rightarrow d;S$
- [3]  $S \rightarrow sT$
- [4]  $T \rightarrow ;sT$
- [5]  $T \rightarrow \varepsilon$

(1) 求文法  $G[P]$  中每个产生式的 **predict** 集。

(2) 构造文法  $G[P]$  的 LL(1) 分析表。

(3) 试应用 LL(1) 分析技术，使用表 1 格式给出句子  $\{d;s;s\}$  的分析过程。

符号栈	输入串	驱动程序动作
#P	{d;s;s}#	
...		

三、设有文法  $G[Z]$  如下：

- $Z \rightarrow AB$
- $A \rightarrow aB \mid b$
- $B \rightarrow aB \mid b$

(1) 构造该文法的 LR(1) 可归活前缀状态机，并构造 LR(1) 分析表。

(2) 试应用 LR(1) 分析技术，使用表 2 格式识别输入符号串 **abb** 是否为文法  $G[Z]$  的句子。

状态栈	符号栈	输入串	分析动作	转向状态
0	#	abb#		
...				

(3)判断该文法是否为 LALR(1)文法，并说明理由。

四、给出如下程序段的四元式中间代码，并利用常表达式节省、公共表达式节省和循环不变式外提三种优化技术对中间代码进行优化。其中 A:array[1..100] of integer，整型变量占 1 个存储单元。

```
a:=0;
j:=(a+1)*5;
while j<100 do
  begin
    a:=a+A[3];
    if a<10 then k:=0;
    else k:=x*y+x*y/5;
    j:=j+1;
  end
```

五、回答下列问题

1.在过程/函数信息表中有一项属性叫做 size，用于存放过程或函数活动记录的大小，请问：

(1)在什么时候填写 size 内容？

(2)在什么时候修改 size 内容？

(3)在什么时候使用 size 内容？

2.我们学习过基于 LL(1)语法分析的语法制导方法，把在语法规则中插入的程序称为语义程序。请问如果采用自顶向下的语法分析方法，且语法分析过程中有回溯的情况，你将采用什么样的措施保证语义程序的正确执行？

3.在生成中间代码的时候，会产生大量的临时变量，请你分析一下临时变量有哪些性质？

## 2014 级编译原理试题

### 一、简答题

1. 高级程序设计语言的实现方式包含哪些？
2. 给出能自动识别只有奇数个 a 和奇数个 b 组成的所有符号串的最简 DFA。
3. 从分析能力和可归前缀自动机状态数目比较一下 LR(0)、SLR(1)、LR(1) 和 LALR(1) 四种语法分析方法。
4. 设数组类型 Array of [1..10] of Array [1..100] of integer，请给出类型的类型表信息。
5. 在语义分析过程中，当处理函数调用语句时，都需要做哪些语义检查？
6. 设有一个类 C 嵌套语言编写的程序如下：

```
void main () {  
    int P(int x)  
    Begin  
        If X > 0 Then  
            Return P(X-1) + X;  
        Else  
            Return 0;  
        End  
    Begin  
        P(100);  
    End  
}
```

请问：当 P 被第 50 次调用而进入函数体后，

- (1) 在内存中为该程序分配的栈区存在多少个活动记录？
- (2) 当前活动记录的局部 Display 表中有几条记录，都记录了什么信息？

### 二、已知文法 G[S] 如下：

$$S \rightarrow aS \mid aL$$
$$L \rightarrow Lb \mid b$$

- (1) 请将文法 G[S] 转化为与之等价的 LL(1) 文法。
- (2) 给出转换后的文法的 LL(1) 分析表。
- (3) 试应用 LL(1) 分析技术，使用表 1 格式分析串 aabb 是否为文法 G[S] 的句子。

符号栈	输入串	动作
#S	aabb#	使用规则？替换
.....	.....	.....

三、设有文法  $G[S]$  如下：

$$S \rightarrow Aa \mid bAe \mid Be \mid bBa$$
$$A \rightarrow d$$
$$B \rightarrow d$$

- (1) 构造该文法的 LR(1) 可归前缀状态机，并构造 LR(1) 分析表。
- (2) 判断该文法是否为 LALR(1) 文法，并说明理由。

四、设有如下程序段，其中  $A$ : Array of  $[1..10]$  of Array  $[1..100]$  of integer，程序中的所有变量都为整型变量，占一个存储单元。

```
x := 0;
s := x + 1;
i := 2 * s;
while i < 10 do
begin
  k := 1;
  while k < 100 do
  begin
    j := 2 * k;
    if j < 50 then
    begin
      s := s + A[3][k] * A[i][k];
    end
    k := 2 * k;
  end
  i := i + 1
end
```

- (1) 给出该程序段代码的四元式中间代码；
- (2) 请利用常表达式节省、公共子表达式节省和循环不变式外提三种优化技术对四元式中间代码进行优化，给出优化后的四元式代码。

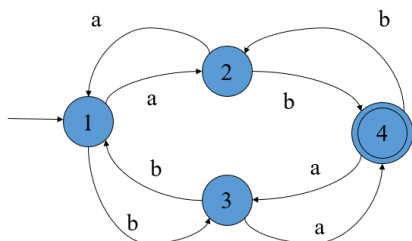
五、高级程序设计语言可以分为嵌套式语言（如 PASCAL 语言）和并列式语言两大类（如 C 语言）两大类，请分析在编译过程中，对于嵌套式语言的处理和并列式语言的处理都有哪些区别？

## 2014 级编译原理试题解答

### 一、简答题

1. 编译实现方式，解释实现方式

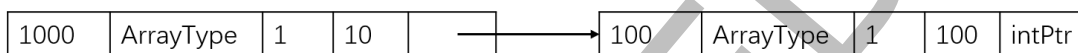
2. 最简 DFA 有四个状态，如下：



3. 分析能力：LR(1) > LALR(1) > SLR(1) > LR(0)

可归活前缀自动机状态数目：LR(0) = SLR(1) = LALR(1) < LR(1)

4.



5. 检查函数是否为函数标识符、检查形参类型与实参类型是否相容、检查形参个数与实参个数是否相等

6. (1) 51

(2) 有两条记录，分别是 main 函数活动记录的首地址和 P 函数活跃活动记录的首地址

二、

(1) 提取公共前缀并消除左递归，得

$S \rightarrow aS'$  (1)

$S' \rightarrow S$  (2)

$S' \rightarrow L$  (3)

$L \rightarrow bL'$  (4)

$L' \rightarrow bL'$  (5)

$L' \rightarrow \epsilon$  (6)

(2) 先求 predict 集，再得出 LL(1) 分析表，如下：

	a	b	#
S	1		
S'	2	3	
L		4	
L'		5	6

(3)

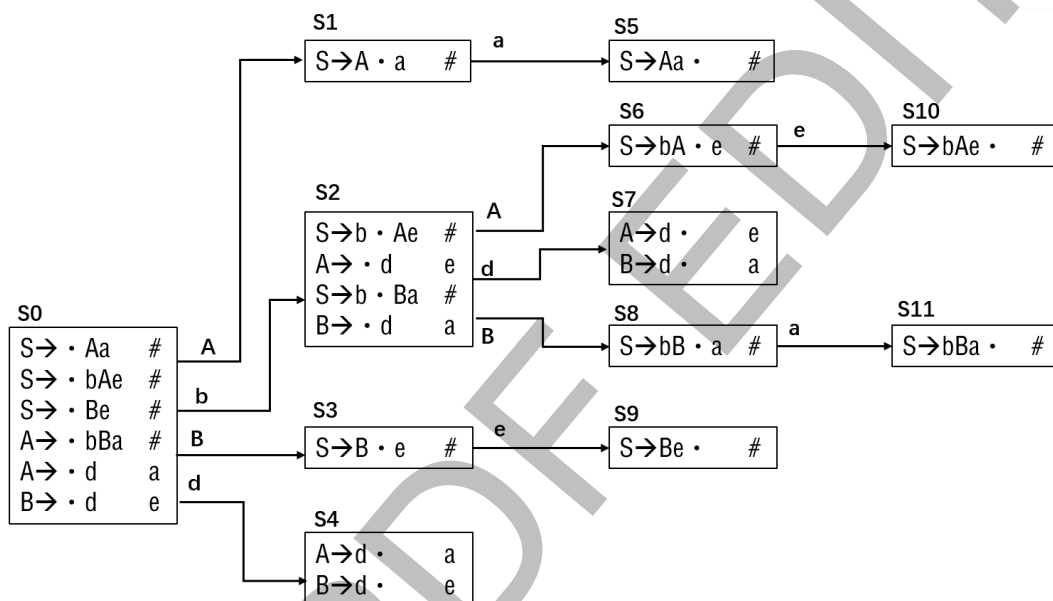
符号栈	输入串	动作
#S	aabb#	使用规则 1 替换
#S' a	aabb#	匹配 a
#S'	abb#	使用规则 2 替换
#S	abb#	使用规则 1 替换
#S' a	abb#	匹配 a
#S'	bb#	使用规则 3 替换

#L	bb#	使用规则 5 替换
#L' b	bb#	匹配 b
#L'	b#	使用规则 5 替换
#L' b	b#	匹配 b
#L'	#	使用规则 6 替换
#	#	成功

三、给文法产生式编号，如下：

- ①  $S \rightarrow Aa$       ②  $S \rightarrow bAe$       ③  $S \rightarrow Be$   
 ④  $S \rightarrow bBa$       ⑤  $A \rightarrow d$       ⑥  $B \rightarrow d$

(1) LR(1)可归前缀状态机如下：



LR(1)分析表如下：

	Action 表			Goto 表	
	a	e	#	A	B
S0					
S1					
S2					
S3					
S4					
S5					
S6					
S7					
S8					
S9					
S10			R2		
S11			R4		

#### 四、中间代码生成与优化

[illegible]



五、

在编译过程中,嵌套式和并列式语言处理差别主要体现在符号表的管理和抽象地址映射两大方面。

(1) 对于符号表的局部化处理,嵌套式语言使用了 Scope 栈,Scope 栈中的 Scope[L]指向 L 层符号表的起始地址;并列式语言可以不使用 Scope 栈,用两个变量分别存储 0 层和 1 层的符号表地址即可。

(2) 分配抽象地址的时候,并列式语言中并列的分程序可以共用

(3) 在抽象地址映射过程中,对于并列式语言,在活动记录可以不为变量访问环境分配空间,使用两个寄存器分别存储静态区和当前活动记录首地址就可以完成抽象地址到物理地址的映射;对于嵌套式语言,在活动记录必须为变量访问环境分配空间。

## 2015 级编译原理试题

### 一、 简答题

1. 简述典型编译器的构成以及各个部分的功能和作用。
2. 设计确定有限状态自动机，识别被 5 整除的二进制正整数（不包括有前导零的数）。
3. 举例说明乔姆斯基文法分类的依据，以及各类文法表示能力的区别。
4. 设计一个上下文无关文法  $G$ ，描述语言  $\{a^i b^j c^k d^l \mid i \geq 0, j \geq 1, k \geq 2\}$ 。
5. 简述语义分析与语法分析的区别，并列举三种语义错误。
6. 设有文法  $G[E]$  如下，求句型  $(E)^* i + i$  的语法树以及所有短语、简单短语和句柄。

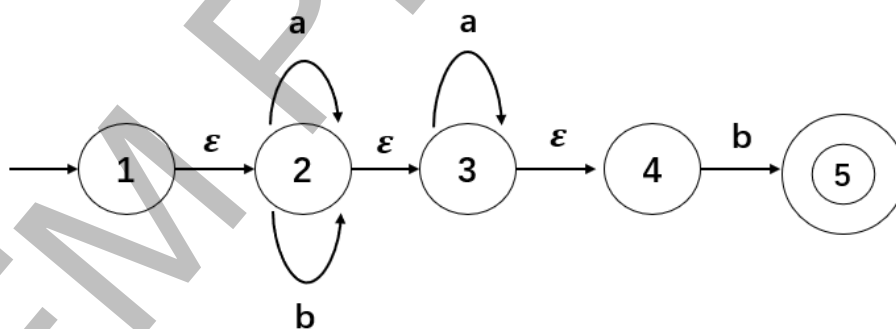
$G[E]: E \rightarrow E+T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow i \mid (E)$

### 二、 问答题

1. 对下图所示非确定有限状态自动机  $M$ ，写出与  $M$  等价的正则表达式，并将  $M$  分别进行确定化和最小化。



2. 有文法  $G[S]$  如下。消除  $G$  的左递归文法；对新文法求每个非终极符的 First 集、Follow 集；构造 LL(1) 分析表；试应用 LL(1) 分析技术，分析输入序列  $abbcd e$  是否为该文法的句子。

$G[S]: S \rightarrow aABe$

$A \rightarrow b \mid Abc$

$B \rightarrow d$

3. 给定 C 语言代码如下所示，使用驻留法对于程序进行语义分析，请画出标识符的符号表内部表示并填写各项存储内容。设函数 f 所在层，形参的起始偏移为  $off_0$ ，整型变量占一个单元，实型变量占两个单元。

```
const float x = 1000;
```

```
int y = 20;
```

```
int f(int x, float y){ return x + 1; }
```

name	kind	type	level	off	access/class	value/param	code	size	forward
------	------	------	-------	-----	--------------	-------------	------	------	---------

### 三、 计算题

1. 已知文法  $G[S]$  的产生式如下，其中 S、A、B 为终极符，\* 为终极符。试回答下列问题并说明理由：

$G[S]: S \rightarrow A*B \mid B$

$A \rightarrow *B \mid *$

$B \rightarrow A$

- 请画出 LR(0) 活前缀自动机，并判断该文法是否为 LR(0) 文法。如果不是有几个状态出现类型冲突；
  - 判断 SLR(1) 方法能否解决这些冲突；
  - 请画出 LR(1) 活前缀自动机，判断该文法是否为 LR(1) 文法。
2. 求以下 C 语言的中间代码表示，以及 main 函数运行到 while 循环中 fab(3) 被调用时内存的分配情况（包括过程活动记录的内部结构和过程活动记录之间的相对位置，形参变量起始偏移为  $off_0$ ）。

```
int fab(int n){
    if(n < 3) return 1;
    else return fab(n-1) + fab(n-2);
}

main(){ int i = 1, n = 10;
    while(i++ < n){
        fab(i); }
}
```

## 2015 级编译原理试题解答

3.根据产生式所受限制的不同，把文法分为四类：

短语型文法： $\alpha \rightarrow \beta$

上下文相关文法： $\alpha A \beta \rightarrow \alpha \gamma \beta$

上下文无关文法： $A \rightarrow \alpha$

线性文法/正则文法： $A \rightarrow \alpha B$  或  $A \rightarrow \alpha$

4.要构造 a,b 指数相同必须有递归，形如  $A \rightarrow aAb$  形式

$S \rightarrow XYZ$

$X \rightarrow aXb \mid \varepsilon$

$Y \rightarrow cY \mid c$

$Z \rightarrow dZe \mid de$

5.语法分析是检查是不是符合语法要求，而语义分析是检查语句的含义是否有误。

If 的条件语句不是 bool 类型

类型赋值错误

变量未声明错误

6.这类题的做法是先构造语法树，再从顶层往下层看。

短语： $(E)^*i+i$ ;  $(E)^*i$ ;  $i$ ;  $(E)$   $i$  有两个，第一个  $i$  和第 2 个  $i$  都是短语

简单短语： $(E)$ ;  $i$  第一个  $i$  和第 2 个  $i$  都是

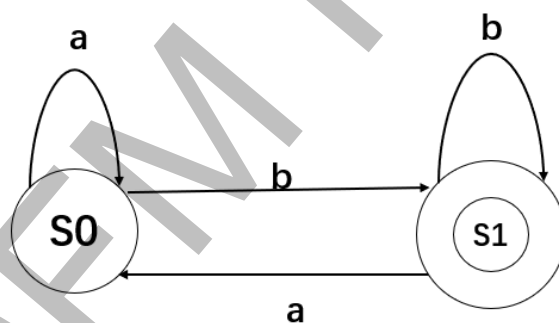
句柄： $(E)$

二、问答题

1.正则表达式： $(a|b)^*a^*b$

	$I_a$	$I_b$
$\{1, 2, 3, 4\}$	$\{2, 3, 4\}$	$\{2, 3, 4, 5\}$
$\{2, 3, 4\}$	$\{2, 3, 4\}$	$\{2, 3, 4, 5\}$
$\{2, 3, 4, 5\}$	$\{2, 3, 4\}$	$\{2, 3, 4, 5\}$

第一个和第二个状态合并后，只有两个状态  $S_0, S_1$



2. $G[S]$ 如下：

$S \rightarrow aABe$

$A \rightarrow bA'$

$A' \rightarrow bcA'$

$A' \rightarrow \varepsilon$

$B \rightarrow d$

于是 First, follow, predict 集如下：

	first	follow
S	{a}	{#}
A	{b}	{d}
A'	{b, ε}	{d}
B	d	{e}

	predict
$S \rightarrow aABe$	{a}
$A \rightarrow bA'$	{b}
$A' \rightarrow bcA'$	{b}
$A' \rightarrow \varepsilon$	{d}
$B \rightarrow d$	{d}

LL(1)分析表如下:

	a	b	c	d	e
S	1				
A		2			
A'		3		4	
B				5	

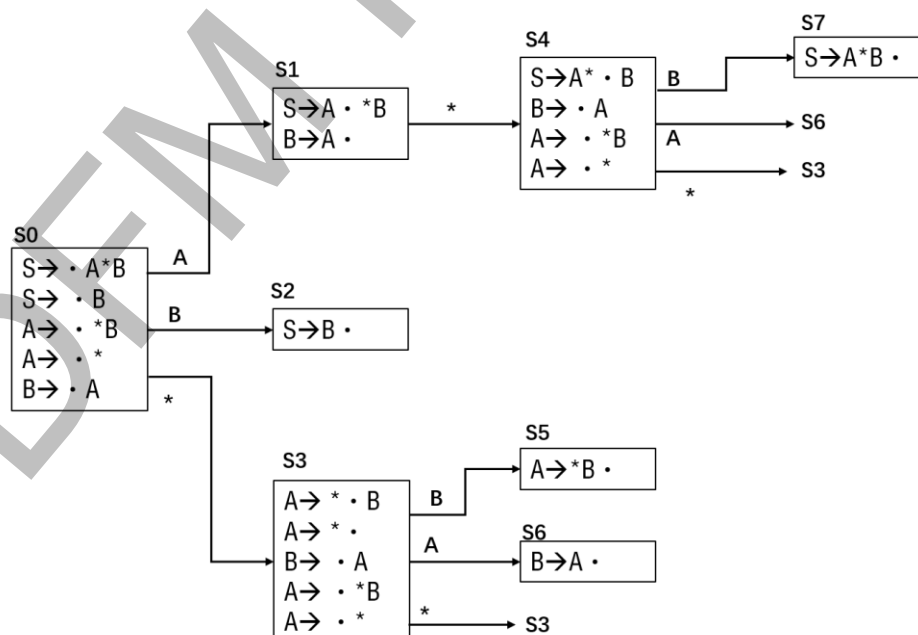
	输入流	动作
S#	abbcde#	使用第一条产生式
aABe#	abbcde#	Match(a)
Abe#	bbcde#	使用第二条产生式
bA'Be#	bbcde#	Match(b)
A'Be#	bcde#	使用第三条产生式
bcA'Be#	bcde#	Match(b);Match(c)
A'Be#	de#	使用第四条产生式
Be#	de#	使用第五条产生式
de#	de#	Match(d);Match(e)

3.如果没优化的话, const float x 应该占空间; 优化的话就不占, 后面有类似题

Name	kind	type	level	off	Access /class	Value/ param	code	size	forward
x	constKind	floatPtr	L	0	dir	1000			
y	varKind	intPtr	L	2	dir	20			
f	routKind	intPtr	L	Off <sub>0</sub>	actual	指向 x	NULL	3	false
x	varKind	intPtr	L+1	Off <sub>0</sub>	dir				
y	varKind	floatPtr	L+1	Off <sub>0</sub> +1	dir				
# 指向第二行									

三、计算题

1. a) LR(0)活前缀自动机如下:

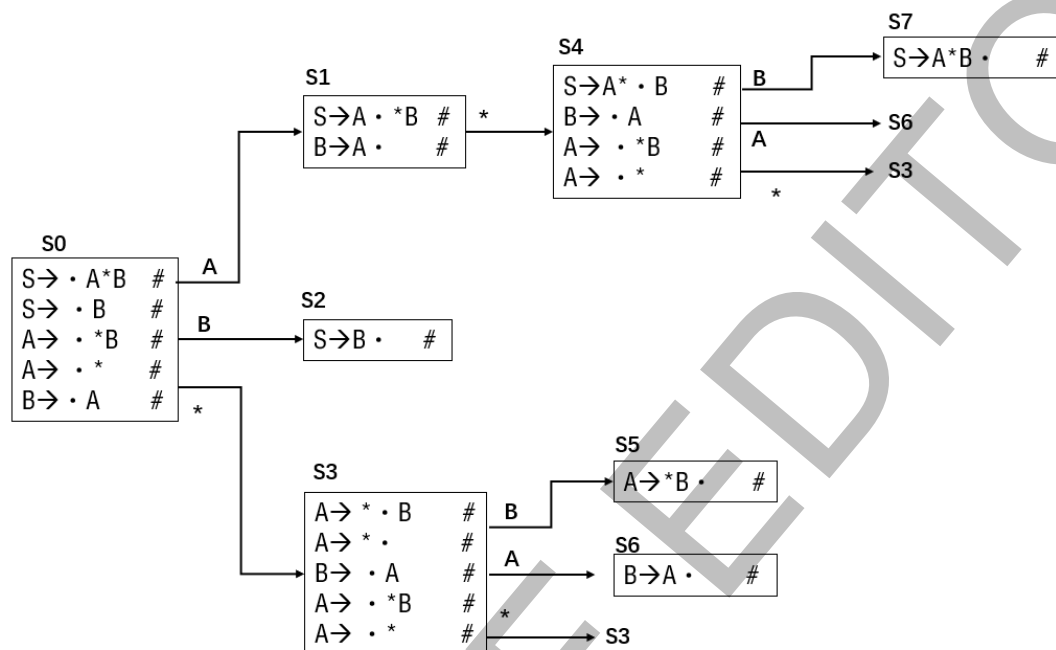


对于 S1 和 S3 状态，有移入-归约冲突，于是不是 LR(0)语法。

b)  $\text{follow}(B) = \{\#, *\}$        $\text{follow}(A) = \{*, \#\}$

由于交集不空，因此对于 S1 和 S3 状态来说都不能向前看一个符号来解决移入-归约冲突

c) LR(1)活前缀自动机如下：



观察可以知道，S1 和 S3 的移入归约冲突还是没用解决，所以该文法不是 LR(1) 文法。

2. 函数的参数传递(ValACT/VarACT, tn, Offset, size)

四元式	解释
(ENTRY, fab, -, -)	Fab 函数
(<, n, 3, t1)	
(THEN, t1, -, -)	
(RETURN, -, -, 1)	Return 1;
(ELSE, -, -, -)	
(-, n, 1, t2)	
(VALACT, t2, off0, 1)	变参传入
(CALL, fab, true, t3)	
(-, n, 2, t4)	
(VALACT, t4, off0, 1)	
(CALL, fab, true, t5)	
(+, t3, t5, t6)	
(RETURN, -, -, t6)	return
(ENDIF, -, -, -)	
(ENDFUNC, -, -, -)	
(ENTRY, main, -, -)	Main 函数
(=, 1, -, i)	

(=, 10, -, n)	
(WHILE, -, -, -)	
(<, i, n, t7)	
(+, i, 1, t8)	I++
(=, t8, -, i)	
(DO, t7, -, -)	
(VALACT, i, off <sub>0</sub> , 1)	
(CALL, fab, true, t9)	Fab(i)
(ENDWHILE, -, -, -)	
(ENDFUNC, -, -, -)	

过程活动记录的内部结构没画

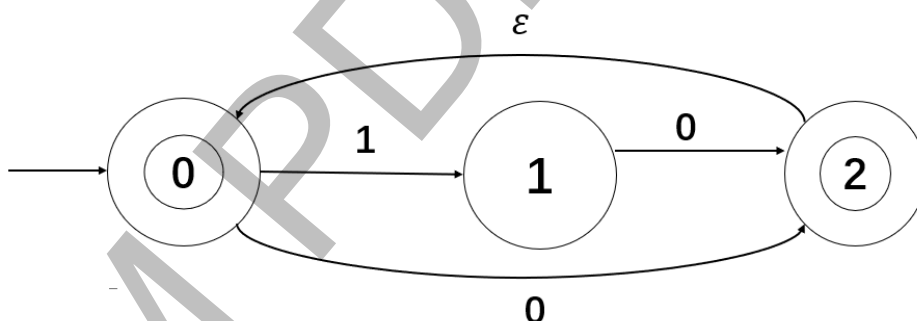
## 2016 级编译原理考题

### 一、简答题（6\*5 分）

1. 请画出编译器的功能结构图。
2. 简述 DFA 和 NFA 有哪些主要的区别？
3. 请画出过程活动记录 AR 应包含哪些内容？
4. 请分布从语法分析能力和状态数上分析以下四种方法的关系：LR(0)、SLR(1)、LR(1)、LALR(1)。
5. 设计一个上下文无关文法 G，可以描述语言  $\{1^n 0^m 1^n \mid n, m \geq 0\}$ 。
6. 设有文法 G[S] 如下，求句型 (S, (a)) 的语法树以及所有短语，简单短语，和句柄。

$G[S]: S \rightarrow (L) \mid aS \mid a$   
 $L \rightarrow L, S \mid S$

### 二、请将下图所示的 NFA 转换为最小 DFA（10 分）



### 三、已知文法 G[S]，请写出它的递归下降 C 语言分析程序（10 分）

$S \rightarrow eT \mid RT$   
 $T \rightarrow DR \mid \epsilon$   
 $R \rightarrow dR \mid \epsilon$   
 $D \rightarrow a \mid bd$



四、对于文法  $G[A]$ ：

$$A \rightarrow aABbc \mid a$$

$$B \rightarrow Bd \mid e$$

试提取  $G[A]$  左公共因子，清除左递归。分析新产生的文法  $G'$  是否为 LL(1) 文法，并构造文法  $G'$  的 LL(1) 分析表（10 分）

五、有下图程序片段，假设程序点的层数和偏移为  $(L, off_0)$ 。试写出各个程序点的层数和偏移的变化情况。（ $d$  为第一个形参可用区距， $int$  占一个单元， $float$  占两个单元）（10 分）

```
① typedef at = int a[10][10];  
② struct { int number; at score; } bt;  
③ at x, *y;  
④ int f(at a, ⑤ at *b, ⑥ float x){  
    int n;  
    .....  
}  
⑧ void g(){  
⑨     int l;  
    .....  
}  
⑩
```

六、已知文法  $G[S]$  的产生式如下，其中  $S$  和  $B$  为非终极符， $a$  和  $b$  为终极符。（15 分）

$$G[S]: S \rightarrow BB$$

$$B \rightarrow aB \mid b$$

试回答下列问题：

- (1) 请画出
- (2) 填写 LL(1) 分析表，并判断该文法是否为 LR(1) 文法；
- (3) 给出符号串  $abb\#$  的分析过程。

七、将下列 C 语言程序片段翻译成四元式（15 分）

```
float A[20];
int i = 0;
float sum = 1.0;
int j = 1;
if(j > 0){
    while(i < 20){
        A[i] = i * 5 - 1;
        sum = sum + A[i] / 2.5;
        i = i + 1;
    }
}
else i = -1;
```

## 2016 级编译原理考题解答

5. 要构造 a, b 指数相同, 必须有  $X \rightarrow aXb$  这种形式, 于是可以得到答案:

$S \rightarrow A \quad A \rightarrow 1A0 \mid B \mid \varepsilon \quad B \rightarrow 0B1 \mid \varepsilon$

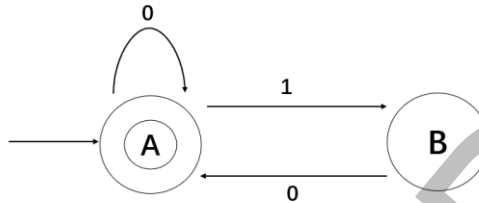
二、

I	I0	I1
$\{0\}^*$	$\{0, 2\}$	$\{1\}$
$\{0, 2\}^*$	$\{0, 2\}$	$\{1\}$
$\{1\}$	$\{0, 2\}$	

有三种状态, 分别设为:  $S_0: \{0\} \quad S_1: \{0, 2\}$

$S_2: \{1\}$

$S_0$  和  $S_1$  是可以合并的, 于是只有两种状态, 设为状态 A, B, 于是最小 DFA 是:



三、先对非终极符号求 first 和 follow 集, 再对文法求 predict 集

	First	Follow
S	$\{e, d, a, b, \varepsilon\}$	$\{\#\}$
T	$\{a, b, \varepsilon\}$	$\{\#\}$
D	$\{a, b\}$	$\{d, \#\}$
R	$\{d, \varepsilon\}$	$\{a, b, \#\}$

文法	Predict 集
$S \rightarrow eT$	$\{e\}$
$S \rightarrow RT$	$\{a, b, d, \#\}$
$T \rightarrow DR$	$\{a, b\}$
$T \rightarrow \varepsilon$	$\{\#\}$
$R \rightarrow dR$	$\{d\}$
$R \rightarrow \varepsilon$	$\{a, b, \#\}$
$D \rightarrow a$	$\{a\}$
$D \rightarrow bd$	$\{b\}$

于是递归下降程序如下:

<pre> Void S(){     If(I ∈ {e}){         Match(a);T();     }else if(I ∈ {d, a, b, #}){         R();T();     }else{         Error();     } }         </pre>	<pre> Void T(){     If(I ∈ {a, b}){         D();R();     }else if(I ∈ {#}){         //no action     }else{         Error();     } }         </pre>
<pre> Void R(){     If(I ∈ {d}){         Match(d);R();     }else if(I ∈ {a, b, #}){         //no action     }else{         Error();     } }         </pre>	<pre> Void D(){     If(I ∈ {a}){         Match(a);     }else if(I ∈ {b}){         Match(b);Match(d);     }else{         Error();     } }         </pre>

四、

$G'[A]$ 如下:

$A \rightarrow aA'$

$A' \rightarrow ABbc \mid \varepsilon$

$B \rightarrow eB'$

$B' \rightarrow dB' \mid \varepsilon$

	first	Follow
A	{a}	{#, e}
A'	{a, $\varepsilon$ }	{#, e}
B	{e}	{b}
B'	{d, $\varepsilon$ }	{b}

文法	Predict 集
$A \rightarrow aA'$	{a}
$A' \rightarrow ABbc$	{a}
$A' \rightarrow \varepsilon$	{#, e}
$B \rightarrow eB'$	{e}
$B' \rightarrow dB'$	{d}
$B' \rightarrow \varepsilon$	{b}

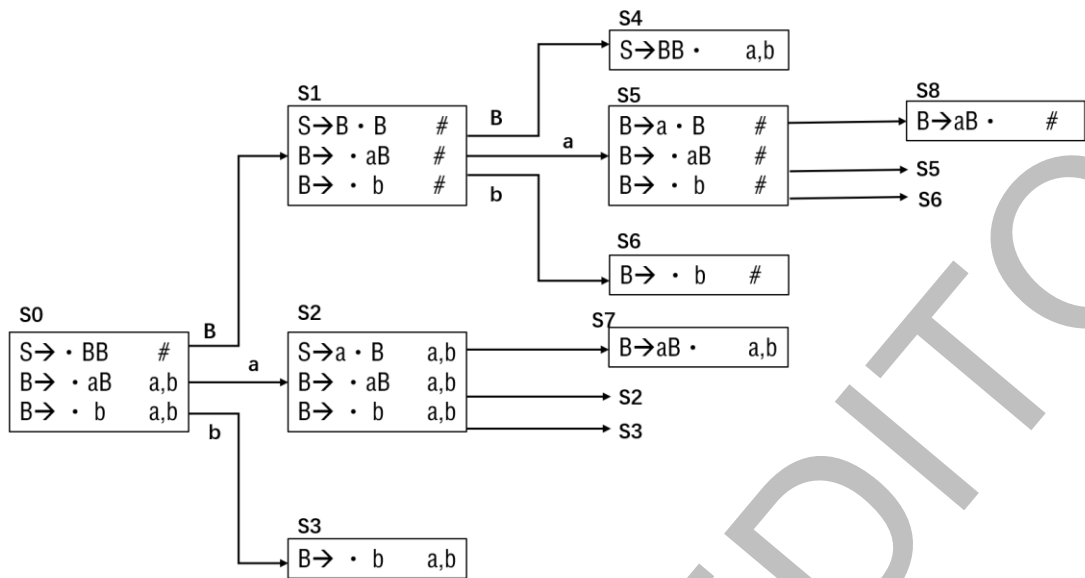
容易看出任何一个非终极符产生式的 predict 集合没有交集, 于是新文法是 LL(1) 文法。其分析表如下:

	a	b	d	e	#
A	1				
A'	2			3	3
B				4	
B'		6	5		

五、类型声明不改变偏移, 只有变量声明改变偏移。

程序点	层数	偏移
2	L	$Off_0$
3	L	$Off_0+101$
4	L	$Off_0+202$
5	L+1	d+100
6	L+1	d+101
7	L+1	d+103
8	L	$Off_0+202$
9	L+1	d
10	L	$Off_0+202$

六、LR(1)活前缀自动机如下：



LR(1)分析表如下，该文法是 LR(1)文法

	action			goto
	a	b	#	B
S0	S2	S3		1
S1	S5	S6		4
S2	S2	S3		7
S3	R3	R3		
S4			Accept	
S5	S5	S6		8
S6			R3	
S7	R2	R2		
S8			R2	

符号串 abb#的分析过程如下：

编号	状态	符号栈	输入流	动作	下一状态
1	0		abb#	S2	2
2	0,2	a	bb#	S3	3
3	0,2,3	ab	b#	R3	7
4	0,2,7	aB	b#	R2	1
5	0,1	B	b#	S6	6
6	0,1,6	Bb	#	R3	4
7	0,1,4	BB	#	Accept	

七、假设 int 占一个单元，float 占两个单元，于是

四元式	解释
(=, 0, -, i)	
(=, 1.0, -, sum)	
(=, 1, -, j)	
(>, j, 0, t1)	
(then, t1, -, -)	If 条件成立
(while, -, -, -)	While 循环标记
(<, i, 20, t2)	
(do, t2, -, -)	
(*, i, 5, t3)	
(-, t3, 1, t4)	
(float, t4, -, t5)	需要一个强制转换四元式子
(-, i, 0, t6)	
(*, t6, 2, t7)	数组一个元素占两个单元
([], a, t7, t8)	A[i]
(=, t5, -, t8)	
(/, t8, 2.5, t9)	
(+, sum, t9, t10)	
(=, t10, -, sum)	
(+, i, 1, t11)	
(=, t11, -, i)	
(ENDWHILE, -, -, -)	循环结束
(ELSE, -, -, -)	
(=, -1, -, i)	
(ENDIF, -, -, -)	If 语句结束

# 2010 级软件学院编译原理试题

选择题

ADBBA

简答题

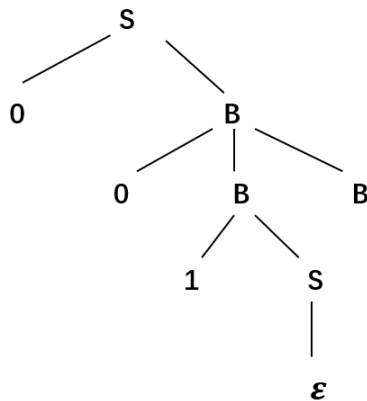
1.

$S \rightarrow AB$

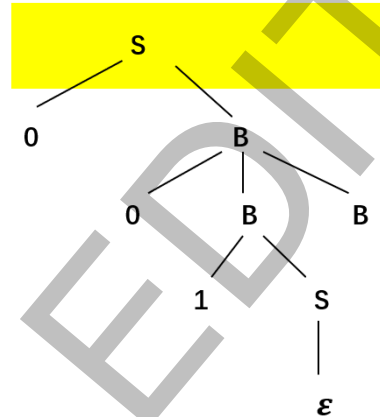
$A \rightarrow aAb \mid ab$

$B \rightarrow bBc \mid bc$

2. 这种类型题，先画语法生成树，然后采取从顶向下观察出短语。



图一



图二

一个快速的方法：先看整体、遮住第一层、遮住第一二层、遮住第一二三层……便可以很快找出短语，简单短语，和句柄

短语：001B 01B 1  $\epsilon$

简单短语： $\epsilon$

句柄： $\epsilon$

3.abdd 不只有一棵语法生成树

4.略

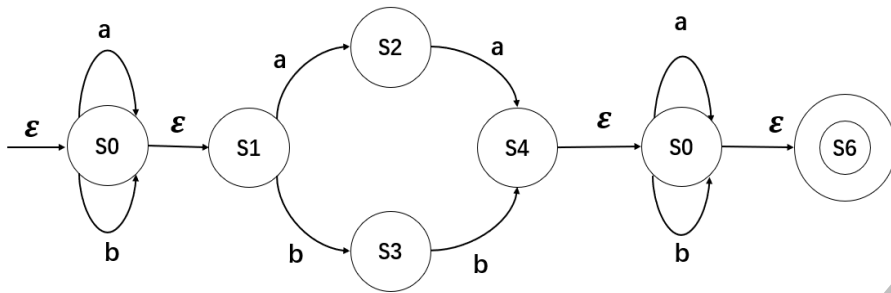
5.输出序列：34242421

6.2013 年试卷相同的考题，略

3.驻留法实现全局顺序表……

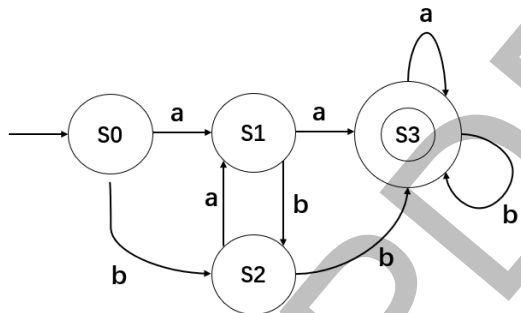
name	kind	type	Access class	level	Offset param	Forward Id
main	routKind	Null	actual	0	Null	false
a	varKind	intPtr	dir	1	initOff	1
b	varKind	intPtr	dir	1	initOff+1	1
b	varKind	intPtr	dir	1	initOff+2	2
a	varKind	intPtr	dir	1	initOff+3	3
#						
b	varKind	intPtr	dir	1	initOff+2	3
#						
#						
#						

4.先画出 NFA，再画表格，再状态合并，最后画出最小的 DFA



等价状态	Ia	Ib
$\{0,1\} +$	$\{0,1,2\}$	$\{0,1,3\}$
$\{0,1,2\}$	$\{0,1,2,4,5,6\}$	$\{0,1,3\}$
$\{0,1,3\}$	$\{0,1,2\}$	$\{0,1,3,4,5,6\}$
$\{0,1,2,4,5,6\}^*$	$\{0,1,2,4,5,6\}$	$\{0,1,3,5,6\}$
$\{0,1,3,4,5,6\}^*$	$\{0,1,2,5,6\}$	$\{0,1,3,4,5,6\}$
$\{0,1,3,5,6\}^*$	$\{0,1,2,5,6\}$	$\{0,1,3,4,5,6\}$
$\{0,1,2,5,6\}^*$	$\{0,1,2,4,5,6\}$	$\{0,1,3,5,6\}$

画表格发现，最后四个等价状态可以合并，于是最小 DFA 如下：



5.先消除左递归，得  $G[S]'$

$S \rightarrow AS'$

$A \rightarrow i := E$

$E' \rightarrow +FE' \mid \varepsilon$

求 first follow predict 三个集合

$S' \rightarrow ;AS' \mid \varepsilon$

$E \rightarrow FE'$

$F \rightarrow (E) \mid i$

	predict
$S \rightarrow AS'$	i
$S' \rightarrow ;AS'$	;
$S' \rightarrow \varepsilon$	#
$A \rightarrow i := E$	i
$E \rightarrow FE'$	( i
$E' \rightarrow +FE'$	+
$E' \rightarrow \varepsilon$	; # )
$F \rightarrow (E)$	(
$F \rightarrow i$	i

	first	follow
S	i	#
S'	; $\varepsilon$	#
A	i	; #
E	( i	; # )
E'	+ $\varepsilon$	; # )
F	( i	+ ; # )



于是 LL(1)分析表如下:

	i	;	#	(	+	)
S	1					
S'		2	3			
A	4					
E	5			5		
E'		7	7		6	7
F	9			8		

6.

$F \rightarrow (\bullet E)$	$\# + * (i$	$T \rightarrow \bullet F$	$(+) i$
$E \rightarrow \bullet E + T$	$) +$	$F \rightarrow \bullet F^*$	$*)+(i$
$E \rightarrow \bullet T$	$) +$	$F \rightarrow \bullet (E)$	$*)+(I$
$T \rightarrow \bullet TF$	$(+) i$	$F \rightarrow \bullet i$	$*)+(i$

输入T	$E \rightarrow T \bullet$	$) +$	$F \rightarrow \bullet F^*$	$* (i$
	$T \rightarrow T \bullet F$	$(i$	$F \rightarrow \bullet (E)$	$(i^*$
			$F \rightarrow \bullet i$	$(i^*$

输入E	$F \rightarrow (E \bullet)$	$\# + * (i$
	$E \rightarrow E \bullet + T$	$) +$

输入F	$T \rightarrow F \bullet$	$(+) i$
	$F \rightarrow F \bullet *$	$*)+(i$

7.要点: @每个 S 对应一个移入动作, 每个 R 对应一个 goto@归约的时候, 要退 n 个状态, 比如归约要用的文法产生式子是  $A \rightarrow abc$ , 则退三个状态。

步骤	状态栈	符号栈	输入流	action	goto
1	0		ababab#	S3	3
2	0,3	a	babab#	R6	2
3	0,2	C	babab#	S4	
4	0,2,4	Cb	abab#	S8	
5	0,2,4,8	Cba	bab#	R7	7
6	0,2,4,7	CbD	bab#	S9	
7	0,2,4,7,9	CbDb	ab#	R5	5
8	0,2,4,5	CbB	ab#	S11	
9	0,2,4,5,11	CbBa	b#	S12	
10	0,2,4,5,11,12	CbBab	#	R3	10
11	0,2,4,5,10	CbBA	#	R1	1
12	0,1	Z	#	Accept	

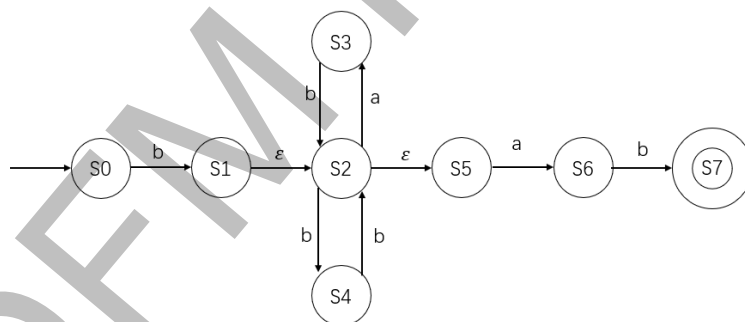
# 2014 级编译原理

中间代码优化的题，肉眼观察法会比常规画表快，但是很容易出错。

编号	优化前四元式	优化后四元式
1	(=, 0, -, x)	(=, 0, -, x)
2	(+, x, 1, t1)	(=, 1, -, s)
3	(=, t1, -, s)	(=, 2, -, i)
4	(*, 2, s, t2)	([], A, 200, t9)
5	(=, t2, -, i)	(WHILE, -, -, -)
6	(WHILE, -, -, -)	(<, i, 10, t3)
7	(<, i, 10, t3)	(DO, t3, -, -)
8	(DO, t3, -, -)	(=, 1, -, k)
9	(=, 1, -, k)	(-, i, 1, t13)
10	(WHILE, -, -, -)	(*, t13, 100, t14)
11	(<, k, 100, t4)	([], A, t14, t15)
12	(DO, t4, -, -)	(WHILE, -, -, -)
13	(*, 2, k, t5)	(<, k, 100, t4)
14	(=, t5, -, j)	(DO, t4, -, -)
15	(<, j, 50, t6)	(*, 2, k, t5)
16	(THEN, t6, -, -)	(=, t5, -, j)
17	(-, 3, 1, t7)	(<, j, 50, t6)
18	(*, t7, 100, t8)	(THEN, t6, -, -)
19	([], A, t8, t9)	(-, k, 1, t10)
20	(-, k, 1, t10)	(*, t10, 1, t11)
21	(*, t10, 1, t11)	([], t9, t11, t12)
22	([], t9, t11, t12)	([], t15, t12, t18)
23	(-, i, 1, t13)	(*, t12, t18, t19)
24	(*, t13, 100, t14)	(+, s, t19, t20)
25	([], A, t14, t15)	(=, t20, -, s)
26	(-, k, 1, t16)	(ENDIF, -, -, -)
27	(*, t16, 1, t17)	(*, 2, k, t21)
28	([], t15, t17, t18)	(=, t21, -, k)
29	(*, t12, t18, t19)	(ENDWHILE, -, -, -)
30	(+, s, t19, t20)	(+, i, 1, t22)
31	(=, t20, -, s)	(=, t22, -, i)
32	(ENDIF, -, -, -)	(ENDWHILE, -, -, -)
33	(*, 2, k, t21)	
34	(=, t21, -, k)	
35	(ENDWHILE, -, -, -)	
36	(+, i, 1, t22)	
37	(=, t22, -, i)	
38	(ENDWHILE, -, -, -)	

8.将下述语句翻译成四元式

编号	四元式
1	(ENTRY, f, -, -)
2	(WHILE, -, -, -)
3	(>, a, 0, t1)
4	(DO, t1, -, -)
5	(+, X, 1, t2)
6	(=, t2, -, X)
7	(<, X, 10, t3)
8	(THEN, t3, -, -)
9	(-, X, 0, t4)
10	(*, t4, 1, t5)
11	([], B, t5, t6)
12	(-, t6, 1, t7)
13	(=, t7, -, a)
14	(ELSE, -, -, -)
15	(-, X, 0, t8)
16	(*, t8, 1, t9)
17	([], B, t9, t10)
18	(/, t10, 2, t11)
19	(=, t11, -, a)
20	(ENDIF, -, -, -)
21	(ENDWHILE, -, -, -)
22	(RETURN, -, -, a)
23	(ENDFUNC, -, -, -)

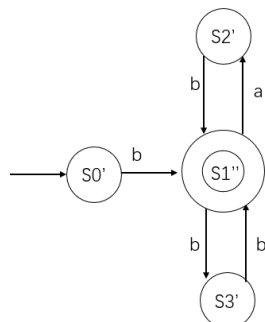
二、请构造与正则表达式  $b(ab|bb)^*ab$  等价的最小 DFA

等效状态	状态集合	Ia	Ib
S0'	{S0} +		{S1, S2, S5}
S1'	{S1, S2, S5}	{S3, S6}	{S4}
S2'	{S3, S6}		{S2, S5, S7}
S3'	{S4}		{S2, S5}
S4'	{S2, S5}	{S3, S6}	{S4}
S5'	{S2, S5, S7} *	{S3, S6}	{S4}

- 1.先画出非确定性自动机
- 2.画表格，来构造确定性自动机
- 3.观察表格，合并状态，观察到  $S1'$ ,  $S4'$ ,  $S5'$  三个状态遇  $Ia$ ,  $Ib$  后下一个状态一样，于是这三个状态可以合并成一个状态  $S1''$ 。

错误的做法，待修正

于是画出最小的 DFA，如下：



五、采用驻留法实现全局线性组织的符号表，给出扫描至下述程序段带下划线语句处时符号表的内容。（**错误答案，待更正**）

Name	Kind	Type	Level	Offset
main	routKind	Null	0	
x	varKind	intPtr	1	0
y	varKind	intPtr	1	1
y	varKind	intPtr	2	2
x	varKind	intPtr	3	3
printf	routKind	Null	3	
x	varKind	intPtr	4	off <sub>0</sub>
y	varKind	intPtr	4	off <sub>0</sub> +1
# 指向黄色的那一行				
# 指向蓝色的那一行				
x	varKind	intPtr	2	2
printf	routKind	Null	2	
x	varKind	intPtr	3	off <sub>0</sub>
y	varKind	intPtr	3	off <sub>0</sub> +1
# 指向蓝色的那一行				
y	varKind	intPtr	2	2
printf	routKind	Null	2	
x	varKind	intPtr	3	off <sub>0</sub>
y	varKind	intPtr	3	off <sub>0</sub> +1
# 指向蓝色的那一行				
printf	routKind	Null	1	
x	varKind	intPtr	2	off <sub>0</sub>
y	varKind	intPtr	2	off <sub>0</sub> +1

## 2013 级唐敖庆班编译原理试题

### 2013 级《编译原理与实现》唐班考试试题(A)

考试时间：2016 年 6 月

班级 \_\_\_\_\_ 学号 \_\_\_\_\_ 姓名 \_\_\_\_\_

- ◇ 请将答案写在答题纸上，写明题号，不必抄题，字迹工整、清晰；
- ◇ 请在答题纸和试题纸上都写上你的班级，学号和姓名，交卷时请将试题纸、答题纸和草纸一并交上来。

#### 一、简答题（每题 5 分，共 40 分）

1. 过程活动记录主要由哪些内容构成？
2. 编译程序按功能分为哪几个阶段？
3. 请给出识别语言  $\{a^{2n+1}b^{n+1} \mid n \geq 0\} \cup \{a^{2n}b^{n+2} \mid n \geq 0\}$  的上下文无关文法，要求不超过 3 条产生式。

4. 写出识别能被 5 整除的非空二进制数串的确有限自动机。

5. 设有如下文法：

$S \rightarrow a \mid (T)$

$T \rightarrow T, S \mid S$

请给出句子  $(a, (a, a))$  的所有短语、简单短语和句柄。

6. 设有如下文法

$A \rightarrow BA \mid \lambda$

$B \rightarrow aB \mid b$

构造该文法的 LR(0) 归约规范活前缀状态。

7. 某语言的文法如下

$P \rightarrow D; S$

$D \rightarrow D; D \mid \text{id}; T \mid \text{proc id}$

$T \rightarrow \text{integer}$

$S \rightarrow S; S \mid \text{id} := n$

其中 D 表示声明，S 表示语句。

写出一个语法制导方案，计算用该语言写的程序中，声明部分声明了多少个 id 及每个 id 在语句部分出现的次数。

8. 假设有如下嵌套的过程声明，若采用全局符号表实现标识符的局部化。请写出在 t4 时刻符号表的主要内容。

```

procedure P()
  var i,j,k:integer; {t1}
  procedure Q()
    var a,b:real;
    procedure R() {t2}
      var a,b:boolean; {t3}
      begin
        R的过程体
      end
    begin
      Q的过程体
    end
  procedure S()
    var c,e:char; {t4}
    begin
      S的过程体
    end {t5}
  begin
    P的过程体
  end

```

二、(15 分)设计一个与正则表达式 $(a|b)^*bb(a|b)^*$ 等价的最小确定有限自动机。

三、(15 分)说明如下文法是否是 LL(1)文法，若不是，将其转换为 LL(1)文法并给出转换后文法的 LL(1)分析表。

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

四、(15 分)已知文法

$Z \rightarrow A$   
 $A \rightarrow BB$   
 $B \rightarrow aB$   
 $B \rightarrow b$

- 1) 构造该文法的 LR(1)归约规范活前缀状态机。
- 2) 构造相应的 LR(1)分析表。

五、(15 分)请写出如下 pascal 程序的局部优化后的四元式形式的中间代码。

```
program main()
  var a:ARRAY[1..10] of integer;
  procedure f(m:integer)
    var b:integer;
    var i:integer;
    begin
      b:=10;
      for i:=1 to 10 do
        a[i]:=m+b*2
      end
    end
  begin
    f(2)
  end.
```

## 2013 年唐敖庆班编译原理试题解答

1. 过程的活动记录主要由哪些部分组成？

动态链指针 返回地址 返回值

寄存器内容 层数 活动记录空间大小 变量访问环境

形参变量区 局部变量区 临时变量区

2. 词法分析 语法分析 语义分析 中间代码生成 中间代码优化 目标代码生成

3. 要求不超过 3 条产生式  $S \rightarrow aaSb$      $S \rightarrow ab$      $S \rightarrow bb$

4. 模 5 的余数有五种情况: 0, 1, 2, 3, 4; 五种状态  $S_0, S_1, S_2, S_3, S_4$

5. 概念: 只含终极符的句型就是句子

短语有:  $(a, (a, a))$      $a, (a, a)$      $(a, a)$      $a$  (三个)    直接短语:  $a$  (三个)    句柄: 第一个  $a$  是

6. 增广文法

$S \rightarrow A$      $A \rightarrow BA | \lambda B \rightarrow aB | b$

7. 假设使用自顶向下方法, 需要四个函数, 两个变量

DCount 代表声明部分 id 出现次数, SCount 代表语句部分 id 出现次数

IntiDCount: 初始化 0

IntiSCount: 初始化 0

AddDCount: DCount 增一

AddSCount: SCount 增一

于是语法制导方案是:

$P \rightarrow D \# \text{IntiDCount} \# ; S \# \text{IntiSCount} \#$

$D \rightarrow D; D \mid id \# \text{AddDCount} \# : T \mid \text{proc } id \# \text{AddDCount} \#$

$T \rightarrow \text{integer}$

$S \rightarrow S; S \mid id \# \text{AddSCount} \# := n$

8. 假设 int 占一个单元, real 占两个单元, boolean 占一个单元, 这题省略 size 是因为过程体有省略, 所以无法求出 size

变量的符号表设计: Name kind type access level offset

函数的符号表设计: Name kind type class level (offset) Param Size Forward

Class  $\rightarrow$  actual | formal (当是 formal 的时候 offset 生效)

Name	Kind	Type	Access class	Level	Offset Param	Forward
P	routKind		actual	0		false
i	varKind	intPtr	dir	1	0	
j	varKind	intPtr	dir	1	1	
k	varKind	intPtr	dir	1	2	
Q	routKind		actual	1		false
a	VarKind	realPtr	dir	2	0	
b	VarKind	realPtr	dir	2	2	
R	routKind		actual	2		false
a	varKind	boolPtr	dir	3	0	
b	varKind	boolPtr	dir	3	1	
# 指向 Q						
S	routKind		actual	1		false
c	varKind	charPtr	dir	2	0	
e	varKind	charPtr	dir	2	1	



二、

1. 画出等价的 NFA

2. 表格法

3. 等效

I	Ia	Ib
$\{S_0, S_1, S_2\}^+$	$\{S_1, S_2\}$	$\{S_1, S_2, S_3\}$
$\{S_1, S_2, S_3\}$	$\{S_1, S_2\}$	$\{S_1, S_2, S_3, S_4, S_5, S_6\}$
$\{S_1, S_2, S_3, S_4, S_5, S_6\}^*$	$\{S_1, S_2, S_5, S_6\}$	$\{S_1, S_2, S_3, S_4, S_5, S_6\}$
$\{S_1, S_2, S_5, S_6\}^*$	$\{S_1, S_2, S_5, S_6\}$	$\{S_1, S_2, S_3, S_5, S_6\}$
$\{S_1, S_2, S_3, S_5, S_6\}^*$	$\{S_1, S_2, S_5, S_6\}$	$\{S_1, S_2, S_3, S_4, S_5, S_6\}$

该题和 15 年唐班试题第一题相同

三、有左递归，还有公共前缀的必不是 LL(1) 文法

$\text{First}(S) = \{ (, a \}$        $\text{First}(L) = \{ (, a \}$

$\text{Follow}(S) = \{ \# \}$        $\text{Follow}(L) = \{ \}, b \}$

$\text{Predict}(S \rightarrow (L)) = \{ ( \}$

$\text{Predict}(S \rightarrow aS) = \{ a \}$

$\text{Predict}(S \rightarrow a) = \{ a \}$

$\text{Predict}(L \rightarrow LbS) = \{ (, a \}$

$\text{Predict}(L \rightarrow S) = \{ (, a \}$

通过 增广、合并、消除左递归

$Z \rightarrow S$

$S \rightarrow (L) \mid aS'$

$S' \rightarrow S \mid \varepsilon$

$L \rightarrow SL'$

$L' \rightarrow bSL' \mid \varepsilon$

文法产生式	Predict 集
$Z \rightarrow S$	$(, a$
$S \rightarrow (L)$	$($
$S \rightarrow aS'$	$a$
$S' \rightarrow S$	$(, a$
$S' \rightarrow \varepsilon$	$b, ), \#$
$L \rightarrow SL'$	$(, a$
$L' \rightarrow bSL$	$b$
$L' \rightarrow \varepsilon$	$)$

文法符号	First 集	Follow 集
Z	$($	$\#$
S	$(, a$	$b, ), \#$
S'	$\varepsilon, (, a$	$b, ), \#$
L	$(, a$	$)$
L'	$b, \varepsilon$	$)$

分析表如下：

	$($	$a$	$b$	$)$	$\#$
Z	1	1			
S	2	3			
S'	4	4	5	5	5
L	6	6			
L'				7	

四、

	First	Follow
Z	a, b	#
A	a, b	#
B	a, b	a, b, #

五、

假定 int 类型占一个单元

四元式

解释

(ENTRY, L1, -, -)

L1: 程序入口

(ENTRY, L2, -, -)

L2: F 函数入口

(=, 10, -, b)

(=, 1, -, i)

For 语句的第一个分号赋值

(FOR, -, -, -)

For 标定循环开始

(<=, i, 10, t1)

条件判断

(DO, t1, -, -)

执行循环

(-, i, 1, t2)

(\*, t2, 1, t3)

([], a, t3, t4)

(\*, b, 2, t5)

(+, m, t5, t6)

(=, t4, -, t6)

(ENDFOR, -, -, -)

(ENDPROC, -, -, -)

(VALACT, 2, 0, 1)

(CALL, f, true, -)

(ENDPROG, -, -, -)

第一个和第二个四元式同样可以写成:

(ENTRY, L1, 16, 1)

(ENTRY, L2, 6, 2)

但是不建议写, 因为涉及到计算, 容易出错

## 2014 级唐班《编译原理与实现》试题 (A 卷)

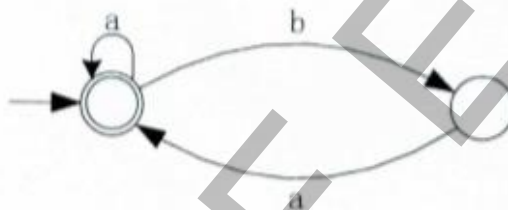
考试时间: 2017 年 6 月

班级 \_\_\_\_\_ 学号 \_\_\_\_\_ 姓名 \_\_\_\_\_

- ◇ 请将答案写在答题纸上, 写明题号, 不必抄题, 字迹工整、清晰;
- ◇ 请在答题纸和试题纸上都写上你的班级, 学号和姓名, 交卷时请将试题纸、答题纸和草纸一并交上来。

### 一、简答题 (每题 5 分, 共 40 分)

#### 1. 给定如下 DFA



请写出与其等价的正则表达式。

2. 构造一个上下文无关文法  $G$ , 使得  $L(G) = \{a^{2m-1}b^m \mid m > 0\}$ 。
3. 在一个移入-归约的分析中采用如下的语法制导方式, 当按某一产生式归约时, 立即执行花括号中的动作, 请问当分析器的输入为  $aacbb$  时, 输出的字符串是什么?

$A \rightarrow aB \{ \text{printf}("0"); \}$

$A \rightarrow c \{ \text{printf}("1"); \}$

$B \rightarrow Ab \{ \text{printf}("2"); \}$

4. 程序运行时, 函数的活动记录是否最多只有一个? 函数的活跃活动记录是否最多只有一个? 回答是或否并作简单说明。

5. 画出如下文法的 LR(0) 归约规范活前缀状态机:

$Z \rightarrow S$

$S \rightarrow S(S) \mid \epsilon$

6. pascal 语言中情形语句结构如下:

CASE E OF

C1: S1;

C2: S2;

... ..

Cn: Sn;

otherwise:  $S_{n+1}$

END

其语义为：若  $E$  的值与  $C_i$  相等则执行  $S_i$ ；若  $E$  与所有  $C_i$  都不相等则执行  $S_{n+1}$ 。请为该情形语句设计一种合理的四元式中间代码结构。

7. 有如下文法：

$A \rightarrow i:=E$

$E \rightarrow E+E$

$E \rightarrow E * E$

$E \rightarrow id$

请问该文法是 SLR(1) 文法吗？为什么？

8. 写出识别被 5 整除的非空二进制数串的最小确定有限自动机。

二、(15 分) 已知文法如下：

$Z \rightarrow S$

$S \rightarrow (SA \mid a$

$A \rightarrow bSA \mid )$

1) 该文法是哪一类 LR 文法？为什么？

2) 构造相应的 LR 分析表。

三、(15 分) 已知文法如下：

$S \rightarrow Aa \mid Ab \mid c$

$A \rightarrow Ad \mid Se \mid f$

1) 将文法等价变换为 LL(1) 文法。

2) 构造变换之后的文法的 LL(1) 分析表。

四、(15 分) 有如下 pascal 示意性源代码，请给出语义分析结束时的驻留法全局符号表（表中应给出标识符的属性信息）。

```
PROGRAM Main
  TYPE at=ARRAY[1..5] of ARRAY[1..10] of real
  VAR b,c: integer;
      d,e:real;
  PROCEDURE p(x:real)
    VAR f:real;
  PROCEDURE q(y:real)
    n:boolean;
    BEGIN IF e<0 THEN p(f) END
  BEGIN q(e) END
  PROCEDURE t(ar1: at; VAR ar2:at)
    VAR j:real;
    BEGIN p(e) END
  BEGIN WHILE e>0 DO p(d) END
```

五、(15 分) 有如下示意性 C 语言程序段，请分别写出该程序段优化前和优化后的四元式中间代码。

```
i=m-1;
j=n;
v=a[n];
while(1)
{
    do i+=1; while(a[i]<v);
    do j-=1; while(a[j]>v);
    if(i>=j) break;
    x=a[i];
    a[i]=a[j];
    a[j]=x;
}
x=a[i];
a[i]=a[n];
a[n]=x;
```

## 2015 级唐敖庆班编译原理试题

- ✧ 请将答案写在答题纸上，写明题号，不必抄题，字迹工整、清晰；
- ✧ 请在答题纸和试题纸上都写上你的班级，学号和姓名，交卷时请将试题纸、答题纸和草纸一并交上来。

一、（20 分）设有限自动机  $M$  接受字母表  $\Sigma=\{0,1\}$  上所有满足如下条件的串：串中至少要包含两个连续的 1。据此回答下列问题：

- (1) 写出与  $M$  等价的正则表达式；
- (2) 构造与  $M$  等价的最小确定有限自动机；
- (3) 构造与  $M$  等价的 3 型文法（正则文法）。

二、（20 分）分支语句 if-then-else 的一个文法如下：

$$S \rightarrow iBtSE \mid a$$
$$E \rightarrow eS \mid \varepsilon$$
$$B \rightarrow b$$

- (1) 计算每一条产生式的 predict 集合；
- (2) 规定“e 与最近的 t 相匹配”，给出相应的 LL(1)分析表。

三、（20 分）根据如下文法，回答问题。

$$S \rightarrow AA$$
$$A \rightarrow Aa \mid a$$

- (1) 构造该文法的 LR(1)归约规范活前缀状态机；
- (2) 根据构造的状态机说明该文法是否为 LR(1)文法，给出理由。

四、（20 分）根据如下 C 程序片段回答问题：

```
int fact(){
    static int i=5;
    if(i==0) return 1;
    else { i=i-1; return (i+abs(1))*fact(); }
}

main()
{ printf("factor of 5=%d\n", fact());}
```

- (1) 写出语义分析结束后，该段程序的全局符号表；
- (2) 写出该段程序的四元式中间代码。

五、(20 分，每题 5 分) 简答 (直接写答案即可)

- (1) 下列 5 项关于过程活动记录的相关描述正确的有哪些？

- ①运行时过程活动记录的管理工作是由编译器生成的目标代码运行时完成的。
- ②若目标机确定，则目标代码生成阶段可以精确计算出每个过程活动记录的大小。
- ③同一个函数被多次调用时所生成的所有过程活动记录的大小一定是相同的。
- ④动态链指针指向栈区、返回地址指向目标代码区。
- ⑤存在间接递归时，递归调用的函数的多个过程活动记录可能不止一个是活跃的。

- (2) “不含公共前缀和左递归的 2 型文法一定是 LL(1)文法”，这种说法是否正确？若正确请证明之，否则举例说明。

- (3) 有如下 C 程序，经过编译连接后，单步跟踪执行时，查看 a 的值，某些调试器会报告“无此变量”。根据所学编译知识解释出现此问题的原因。

```
main(){ int x;  
        const int a=1;  
        x=100+a;  
        printf(“%d\n”, x);}
```

- (4) 表达式 “E1 op E2” 的代码生成顺序采用如下策略：若 E1 和 E2 中都有函数调用，则按从左向右顺序生成；若 E1 和 E2 中只有一个含有函数调用，则先生成含有函数调用的那个子表达式。在该策略下，若将题四中的表达式 “(i+abs(1))\*fact()” 替换为 “(i+1)\*fact()”，其余程序不变 (abs 为求绝对值库函数)，则程序的执行结果会不同，请写出替换前和替换后程序的输出结果。

## 2015 级唐敖庆班编译原理试题解答

1. 注意第三问要求的是 3 型文法

1)  $(0|1)^*11(0|1)^*$

2)

2.

	first	follow
S	{i, a}	{e, #}
E	{e, $\varepsilon$ }	{e, #}
B	{b}	{t}

	predict
$S \rightarrow iBtSE$	i
$S \rightarrow a$	a
$E \rightarrow eS$	e
$E \rightarrow \varepsilon$	e, #
$B \rightarrow b$	b

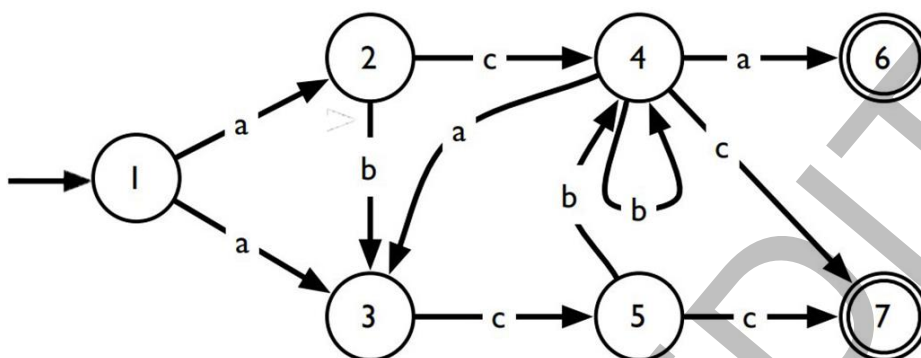
由 predict 集合看出，其不符合 LL(1)文法，但是可以定义规则消除冲突，规定“e 与最近的 t 相匹配”，那么 LL(1)分析表如下：

	i	a	b	e	#
S	1	2			
E				3	4
B			5		



## 2016 级唐敖庆班考试题

一、（10 分）设有如下 NFA:



(1) 构造与之等价的 DFA;

(2) 完成自动机的最小化。

二、（10 分）一个机器人在一个平面空间移动，已知机器人在接到命令 F 时可以向前移动，接到命令 L 时左转 90 度，接到命令 R 时右转 90 度。请给出一个文法 G，该文法 G 的语言为：使得机器人的最终方向与初始时方向相同的所有命令序列（例如 FLR、RRRR）。

三、（30 分）考虑有如下文法:

$G: S \rightarrow a \mid SSS \mid S@S$

(1) 判断该文法是否为二义性文法并说明理由。（5 分）

(2) 假设存在以下规则:

a) 运算符 \$ 是右结合、低优先;

b) 运算符 @ 是左结合高优先。

给出一个新的文法 G1 使得其描述与 G 相同的语言，并符合上述规则。（5 分）

(3) 构建 G 的 LR(0) 可归前缀状态机（需添加拓广产生式）。（10 分）

(4) 分别求出 First(S) 和 Follow(S)，并结合问题 3 中的自动机回答和解释在该 LR(0) 可归前缀状态机中，哪些冲突不能被 LR(0) 语法分析方法解决但可以被 SLR(1) 方法解决，哪些冲突不能被 SLR(1) 方法解决。（10 分）

四、(10 分) 设有如下程序:

```
const m=333;
```

```
const n=444;
```

```
Type at=array[1..10] of real; //类型声明, 长度为 10 的实型数组
```

```
rt=record i,j:integer end; //类型声明, 含有整形 i,j 的结构体
```

```
Var a,b:at; x,y:real; //变量声明
```

已知在分步执行程序进行 debug 时, 若试图查看 m 值时会得到提示“无此变量”, 且设程序当前层数为 L, 偏移为 Offset, 整形占 1 个空间, 实型占 2 个空间。写出本层符号表的内容。

五、(20 分) 有关于变量的文法如下:

$V \rightarrow id$

$V \rightarrow V[E]$

当采用语法制导的方法实现语义分析及中间代码生成时, 应该在文法中增加什么样的语义动作以及这些语义动作要实现什么功能才能实现:

(1) 语义错误检查 (最少写出 3 种);

(2) 生成含有数组越界检查的中间代码, 并请解释为什么一般的编译器中没有数组越界检查。(数组越界是一种目标程序运行时可能发生的 runtime error。例如声明了一个数组为 `int a[100]`, 程序中出现 `a[i]`, 且在程序运行时 `a[i]` 中的 i 在某一时刻为 100, 就发生了数组越界)。

六、(10 分) 如果语言中允许在嵌套声明的内层函数体中用 `goto L` 语句跳转到外层函数的函数体, 即允许非正常出口, 若 L 位于 Q 的函数体中,

(1) 跳出后的新动态 AR 链应怎么变化, 并举例说明?

(2) 编译程序需要额外保存哪些信息, 并如何求得新的 SP 和 TOP 的值?

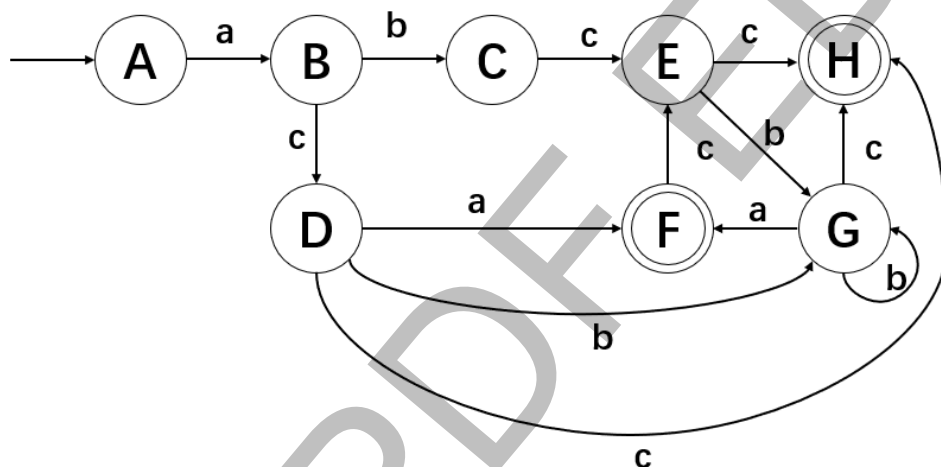
## 2016 级唐敖庆班编译原理试题解答

一、解析：NFA 化成等价的 DFA，先画表；要求最简的 DFA 则需要状态合并。

第一步：画表

		Ia	Ib	Ic
<b>A</b>	$\{1\}^+$	$\{2,3\}$		
<b>B</b>	$\{2,3\}$		$\{3\}$	$\{4,5\}$
<b>C</b>	$\{3\}$			$\{5\}$
<b>D</b>	$\{4,5\}$	$\{3,6\}$	$\{4\}$	$\{7\}$
<b>E</b>	$\{5\}$		$\{4\}$	$\{7\}$
<b>F</b>	$\{3,6\}^*$			$\{5\}$
<b>G</b>	$\{4\}$	$\{3,6\}$	$\{4\}$	$\{7\}$
<b>H</b>	$\{7\}$			

于是等价的 DFA 如下：

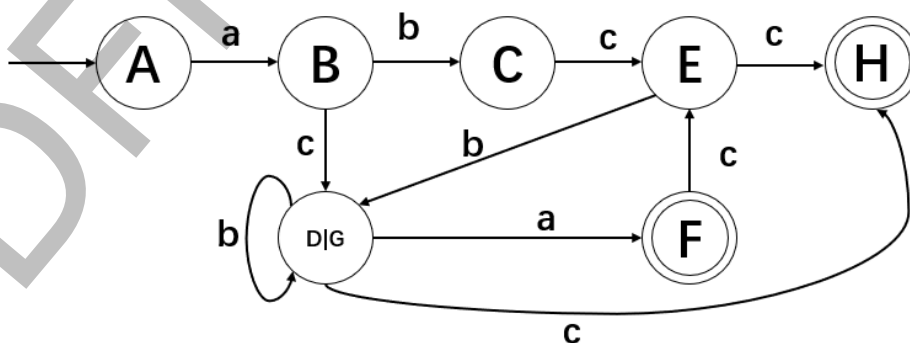


第二步：最简化，采用分割法

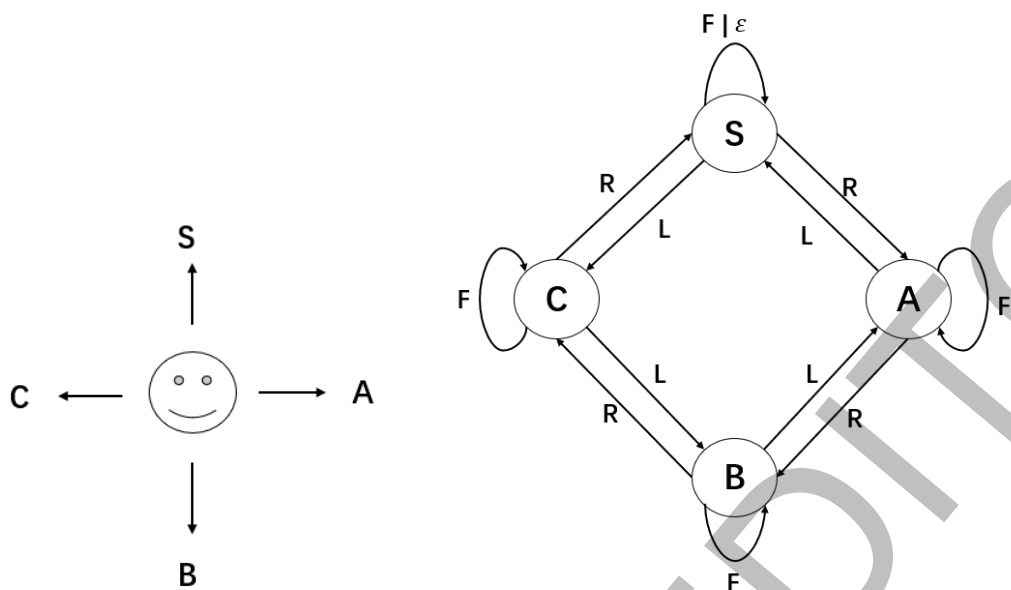
非终止集： $\{A,B,C,D,E,G\} \rightarrow \{A\} \cup \{B\} \cup \{C\} \cup \{D,G\} \cup \{E\}$

终止集： $\{F,H\} \rightarrow \{F\} \cup \{H\}$

于是最简的 DFA 如下：



二、解析：该题可以先画状态转换图，然后写文法；但是如果对文法熟练，可以直接写文法。



答：设 S 是向前状态；A 是朝向右状态；B 是朝向后状态；C 是朝向左状态；R 是向右转命令；L 是向左转命令；F 是向前转命令，于是文法 G 如下：

$S \rightarrow FS   \epsilon$	$S \rightarrow RA$	$S \rightarrow LC$
$A \rightarrow LS$	$A \rightarrow FA$	$A \rightarrow RB$
$B \rightarrow LA$	$B \rightarrow FB$	$B \rightarrow RC$
$C \rightarrow LB$	$C \rightarrow FC$	$C \rightarrow RS$

三、解析：

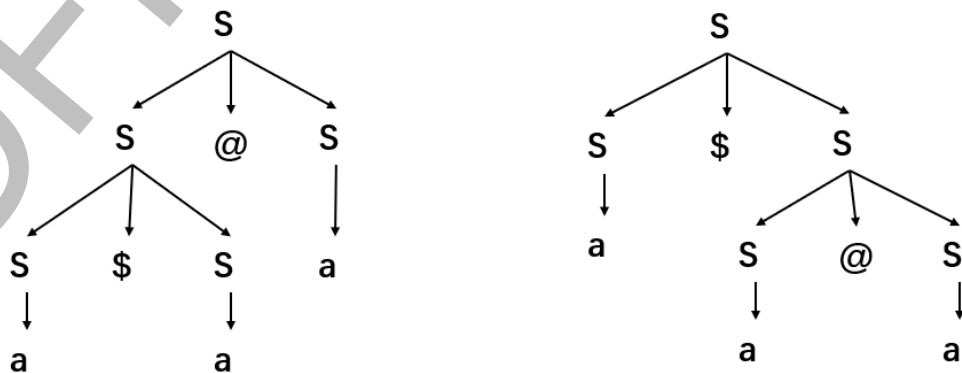
(1) 要判断一个文法有二义性，画出两个不同语法生成树即可；判断一个文法一定没用二义性，比较难，记住一个结论：满足 LL(1) 语法的文法一定没有二义性。

(2) 如果有很强的散发思维的话，可以把 @ 看成乘法，\$ 看成是加法，于是文法 G1 可以写出来（教材上有，但是没有考虑左结合、右结合）

(4) 什么是 LR(0) 冲突呢？就是这个状态存在可归约又存在可移入的项；如果一个状态的归约项的 follow 集合与其可移入的字符集合交集不为空，那么就不符合 SLR(1) 语法。

答：

(1) a\$a@a 有两个不同的语法生成树



(2) 参考加法和乘法的文法产生式，修改即可

$E \rightarrow T$  (1)

$E \rightarrow T\$E$  (2)

$T \rightarrow F$  (3)

$T \rightarrow T@F$  (4)

$F \rightarrow a$  (5)

第二条产生式文法是右结合，第四条满足左结合

(3) 增加拓广产生式后，文法如下：

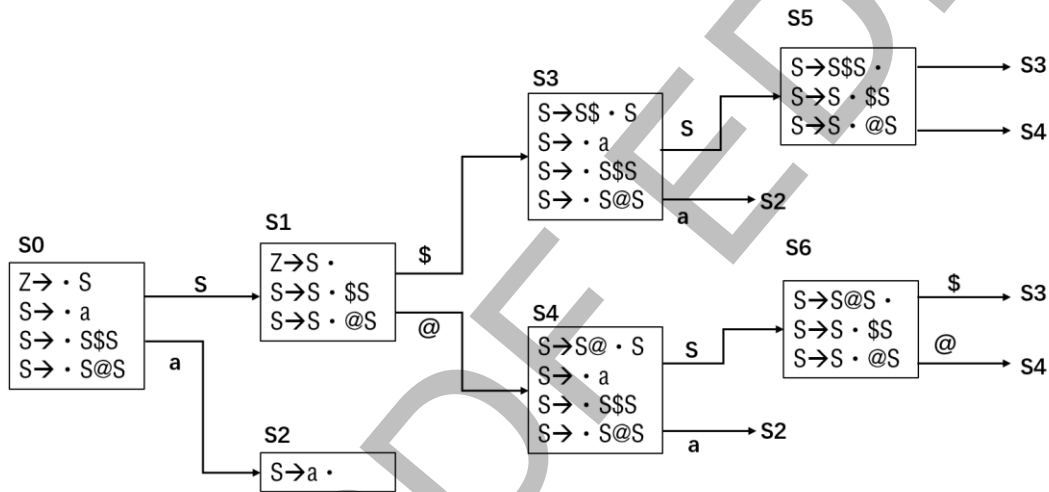
$Z \rightarrow S$

$S \rightarrow a$

$S \rightarrow S\$S$

$S \rightarrow S@S$

于是 LR(0) 可归前缀状态机如下：



(4)  $\text{first}(S) = \{a\}$        $\text{follow}(S) = \{\#, \$, @\}$

有 LR(0) 冲突的状态有：S1, S5, S6；SLR(1) 可以解决 S1 的冲突，但是不能解决 S5 和 S6 的冲突。

四、解析：为什么会无此变量呢？原因是因为程序编译优化了，想想四元式优化中的常量表达式优化，常量它就直接替换了，所以 debug 当然提示无此变量。于是本层符号表的常量不分配偏移即可。

答：本层符号表的内容如下：

Name	Kind	Type	Access	level	Offset	Value
m	constKind	intPtr	dir	L		↑ (333)
n	constkind	intPtr	dir	L		↑ (444)
a	varKind	at	indir	L	Offset	Null
b	varKind	at	indir	L	Offset+20	Null
x	varKind	realPtr	dir	L	Offset+42	Null
y	varKind	realPtr	dir	L	Offset+42	Null

Name	Kind	Type
at	typeKind	ptr

ptr →	10	arrayType	1	10	realPtr
-------	----	-----------	---	----	---------

五、解析：（1）要求是在这里添加语义动作，以及要说明这个动作要实现什么功能，数组的语义错误检查有哪些呢？从最基本的想，首先这个变量 `id` 是否是已经声明？其次，要检查 `V[E]` 是不是数组类型，还要检查数组下标 `E` 是不是合法。

（2）检查数组越界的中间代码就是与上下界做比较

解答：（1）

$V \rightarrow id \text{ \#Live\#}$

$V \rightarrow V[E \text{ \#CheckMark\#}] \text{ \#CheckArray\#}$

各动作符的含义如下：

**Live**：用来检查有没有这个变量 `id` 声明

**CheckMark**：检查 `E` 是不是一个数组下标类型

**CheckArray**：检查 `V[E]` 是不是一个数组类型，比如 `a[n]`，变量 `a` 确实声明了，但是它是一个类型定义类型的，显然不对，是一个语法错误。

（2）需要加的动作是：

$V \rightarrow V[E] \text{ \#OutBound\#}$

数组越界的中间代码就是检查当前的数组下标是不是在上界 `U` 和下界 `L` 之间，如果是，则没有越界，如果不是则越界，转错误处理。中间代码略。

没有数组检查的原因是：可以提高目标代码效率，因为出现数组基本上会与循环一起用，每次循环都检查数组下标是否越界，会多执行很多条指令，导致目标程序效率很低，所以现在的语言编译时基本上不检查数组是否越界。

六、解析：允许 `goto` 语句的话，也就是允许退出多个过程活动记录。这题有待增加更加详细解释，个人能力有限，解释不清。

```
Void P{
    Void Q{
        .....
        L: int a;
        int b;
        Void R{
            int c;
            goto L;
        }
    }
}
```

解答：

（1）找到当前活跃的过程活动记录 `Q`，把该记录后面的记录删除

（2）知道 `L` 所在的层数，这样就知道了 `Q` 的层数，也就能够找到 `Q` 函数所在当前活跃过程活动记录的 `sp`，这个就是新的 `sp` 的值，知道了 `sp`，那么就可以求新的 `top`，`top`=当前 `sp`+过程活动记录大小（`size`）。

## 2017 级唐敖庆班编译原理考试题

一、多选题（4 分\*5）规则：错选得 0 分，少选一个得 2 分，少选超过一个得一分，不选得 0 分

1. 下列哪些字符串可以由正则文法  $((a|\epsilon)b^*)^*$  推导而出？

- A.  $\epsilon$
- B. abb
- C. abab
- D. aaabbbbbbb

2. 下列哪些文法是无二义性文法？

- A.  $S \rightarrow *SSS \mid +SS \mid a$
- B.  $S \rightarrow +S*S \mid *S+S \mid a$
- C.  $S \rightarrow +SS \mid *SS \mid \epsilon$
- D.

3. 下列哪些文法产生式可以推出字符串 aaaaaaa？

- A.  $S \rightarrow SSS \mid a$
- B.  $S \rightarrow SSSS \mid SSS \mid a$
- C.  $S \rightarrow AB; A \rightarrow Aa \mid a; B \rightarrow b \mid \epsilon$
- D.  $S \rightarrow aSa \mid \epsilon$

4. 关于下面的陈述，错误的是？

- A. 存在间接递归时，递归调用的函数的多个过程活动记录可能不止一个是活跃的
  - B. 目标代码(ADD, R, x), 其中 x 代表的是 x 的地址
  - C. 文法  $G: E \rightarrow E*T \mid T, T \rightarrow T+F \mid F, F \rightarrow i$ , 则表达式  $1+2*8+6$  的归约结果是 42
  - D. 略
5. 中间代码生成的临时变量的性质有哪些？
- A. 如果不优化，它只定义和使用一次
  - B. 临时变量是基本数据类型
  - C. 略
  - D. 略

二、简答题（6 分\*5）

1. 画出能被 3 整除的十进制数的 DFA。注意：不允许出现前导 0，即 099 是非法的，不需要考虑负数。

**一定要注意，不允许出现前导 0**

2. 从状态数和描述能力来对比 LR(0)、SLR(1)、LR(1) 和 LALR(1)？

3. 什么是过程活动记录？它包含哪些内容？

4. 在过程/函数信息表中有一项属性叫做 size，用于存放过程或函数活动记录的大小，请问：

(1) 在什么时候填写 size 内容？

**中间代码生成时候**

(2) 在什么时候修改 size 内容？

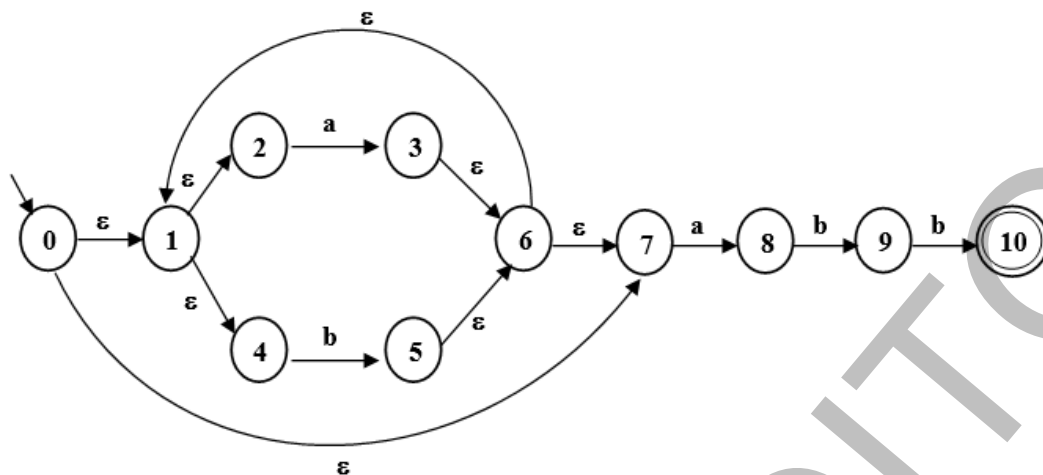
**中间代码每生成临时变量的时候**

(3) 在什么时候使用 size 内容？

**求 top 的时候，用  $sp+size$**

5. 在处理函数调用时，需要检查什么语义信息？

三、将下面的 NFA 确定化（10 分）



四、已知文法 G 如下所示，请构造 LL(1) 分析表（10 分）

$S \rightarrow (L) \mid a$

$L \rightarrow L=S \mid S$

五、有如下文法（10 分）

$Z \rightarrow eBa \mid eAb \mid fBb \mid fAa$

$A \rightarrow h$

$B \rightarrow h$

- (1) 给出 LR(1) 归约活前缀状态图；
- (2) 它是 LALR(1) 文法吗？请说明理由。

六、Display 表

一个类 C 语言的递归的程序片段如下，采用的是局部 display 表方法，问第二次调用递归的时候，栈区内容。

```

void g(){
    int i, j;
    void f(int n){
        ...
    }
    f(10);
}
void main(){
    int a, b;
    g();
}
  
```

七、语法制导

变量声明的语法制导，要求给出对应动作的具体实现，实际上就是填符号表。