

# 计算机与控制工程学院、软件学院 2016—2017 学年 第一学期本科生编译系统原理期末考试试卷(A 卷)

专业：\_\_\_\_\_ 年级：\_\_\_\_\_ 学号：\_\_\_\_\_

姓名：\_\_\_\_\_ 成绩：\_\_\_\_\_

得 分

## 一、 单项选择题（每空 2 分，共 24 分）

1. 将字符流切分为单词流是在 A 阶段，生成三地址码程序是在 D 阶段，对汇编码进行窥孔优化是在 F 阶段。  
A. 词法分析                      B. 语法分析  
C. 语义分析                      D. 中间代码生成  
E. 代码优化                      F. 目标代码生成
2. GCC 主要运行在 Linux 平台，支持 C、C++、Fortran 等多种语言，为此，实现编译器最好的方式是 C。  
A. 单前端单后端                  B. 单前端多后端  
C. 多前端单后端                  D. 多前端多后端
3. 关于正则表达式，下面说法 不正确 的是 D。  
A. | 运算满足交换律              B. 可以用表达式树描述  
C. 可描述活前缀集合              D. 闭包运算满足交换律
4. 下面的正则表达式和上下文无关文法，定义了 A 的语言。  
 $1^*(0 \mid 01)^*$                        $S \rightarrow 0A \mid 1S \mid \varepsilon$   
    $A \rightarrow 0A \mid 1B \mid \varepsilon$   
    $B \rightarrow 0A \mid \varepsilon$   
A. 相同                              B. 不同
5. 关于下面上下文无关文法，说法 不正确 的是 C。  
 $S \rightarrow aAd \mid bBd \mid aBe \mid bAe$   
 $A \rightarrow c$   
 $B \rightarrow c$   
A. 存在等价的正则表达式  
B. 存在等价的 3 型文法  
C. 是 SLR(1) 文法  
D. 是 LR(1) 文法

6. 对预测分析法、算符优先分析算法、SLR 分析算法，下列说法正确的是\_\_C\_\_。

- A. 它们都不能分析左递归文法
- B. 它们都不能分析有左公因子的文法
- C. 它们都不能直接分析二义性文法
- D. 它们都能分析所有定义了正规集的文法

7. 对下面属性定义，说法正确的是\_\_B\_\_。

$A \rightarrow QR \{ R.i = f_1(A.i); Q.i = f_2(R.s); A.s = f_3(Q.s) \}$

$Q \rightarrow a \{ Q.s = f_4(a.v, Q.i); \}$

$R \rightarrow b \{ R.s = f_5(b.v, R.i); \}$

- A. 只包含综合属性
- B. 可编写算法遍历语法树实现翻译
- C. 可与预测分析法结合进行翻译
- D. 可与 LR 分析法结合进行翻译

8. 下面哪种语言有运行时数组下标越界检查机制？\_\_A\_\_。

- A. Ada
- B. BASIC
- C. C++
- D. PASCAL

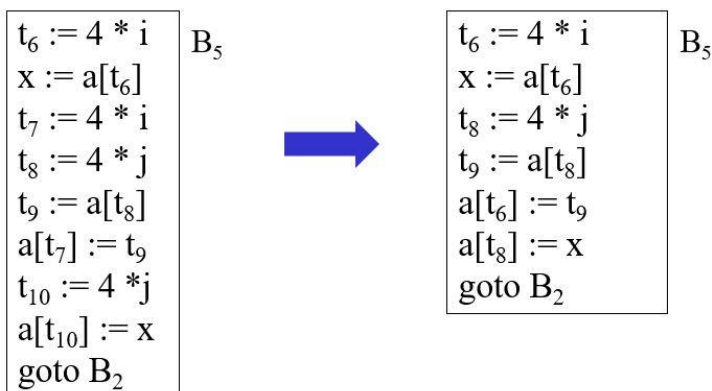
9. 对下面两条汇编指令，说法正确的是\_\_B\_\_。

(1) MOV R0, a

(2) MOV a, R0

- A. (2)肯定是无用指令，可删除
- B. (2)不一定无用，要结合程序其他部分看

10. 下面的基本块内优化，采用了\_\_A\_\_。



- A. 消去公共子表达式
- B. 复制传播
- C. 代码外提
- D. 强度削弱

得 分

二、 设计题（每题 6 分，共 24 分）

1. 一个数据集中每行是一条记录，一条记录由 3 个字段组成，字段间空格分隔，第二个字段是 0~1（不包括 1）的两位小数，另两个字段是 0~255（包括 255）间的整数，数值都以文本方式保存，例如某个字段是 25 0.70 164。设计正则表达式描述数据集。

答：

num->[0-9] | [1-9][0-9] | 1[0-9][0-9] | 2[0-4][0-9] | 25[0-5]

dataset->(num' '\*0'.'[0-9][0-9]' '\*num)+

或等价正则表达式

**\*字段 1、3 错一个扣 1 分，错两个 3 分；字段 2 和最后的闭包错 1 个扣 1 分，错两个扣 3 分**

2. 设计正则表达式，接受能被 4 整除的八进制数（八进制数以一个前导 0 表示）。

答：

00 | 04 | 0[1-7][0-7]\*(0|4)

或等价正则表达式

**\*漏掉 00、04 扣 1 分，最后一种情况，漏掉前导 0 扣 1 分，(0|4)错了扣 1 分，[1-7][0-7]\*错了扣 2 分**

3. 设计上下文无关文法描述 Pascal 语言的普通变量声明语句，类型可以是 integer、char、real、boolean，形如 a, b, c : integer。

答：

D-> L : T

T-> integer | char | real | boolean

L->L, id | id

或等价 CFG

**\*每个产生式错了扣 2 分**

4. 给出函数 find\_char 的类型表达式。

struct pos { int first; int last; };

pos find\_char(char \*str, char c);

答：

pos 的类型表达式 record(int×int)

find\_char 的类型表达式 pointer(char)×char-> record(int×int)

\*pointer(char)×char 3 分， record(int×int) 3 分

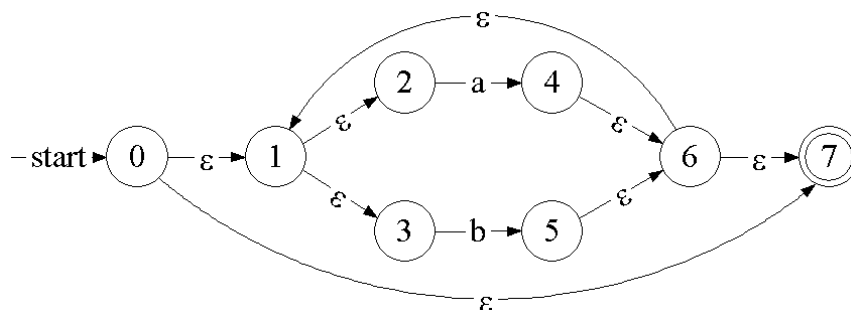
得分

三、（22 分）对下面的正则表达式。

$(c | d)^*$

1. 用 Thompson 构造法将其转换为 NFA。（7 分）

答：其中 a 改为 c，b 改为 d



\*按正确的比例酌情给分

2. 用子集构造法将得到的 NFA 转换为 DFA，写出识别 ababbab 的状态转换序列。（9 分）

答：

$\epsilon\text{-closure}(\{0\}) = \{0, 1, 2, 3, 7\} = A$

$\epsilon\text{-closure}(\delta(A, a)) = \{1, 2, 3, 4, 6, 7\} = B$

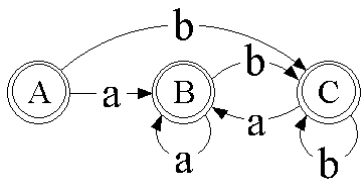
$\epsilon\text{-closure}(\delta(A, b)) = \{1, 2, 3, 5, 6, 7\} = C$

$\epsilon\text{-closure}(\delta(B, a)) = B$

$\epsilon\text{-closure}(\delta(B, b)) = C$

$\epsilon\text{-closure}(\delta(C, a)) = B$

$\epsilon\text{-closure}(\delta(C, b)) = C$



ababbab 状态转换序列：

$A \rightarrow B \rightarrow C \rightarrow B \rightarrow C \rightarrow C \rightarrow B \rightarrow C$

\*计算过程 4 分，状态转换图 3 分，状态转换序列 2 分

3. 将 DFA 最小化（6 分）

答：

初始划分  $\{A, B, C\}$

对符号 a 的状态迁移都转到 B，对符号 b 的状态迁移都转到 C  
因此不能继续分裂，最终结果为

状态集{A}

字母表{a, b}

初态 A

终态集{A}

状态迁移函数 $\delta(A, a) = \{A\}$ ,  $\delta(A, b) = \{A\}$

**\*计算过程 3 分，状态转换图（一个状态既是初态又是终态，两条边都指向自身） 3 分**

得 分

四、（16 分）对下面前缀表达式文法

$E \rightarrow + E E \mid - E E \mid * E E \mid / E E \mid \text{num}$

（1）列出终结符、非终结符和开始符号（3 分）

（2）文法中没有描述带括号的表达式，是否不完整？为什么？

（3 分）

（3）设计**语法制导定义**将其转换为中缀表达式（注意保持正确的运算顺序），画出\*9+52 的语法树，计算属性值将其扩展为注释语法树（10 分）

答：

（1）终结符：+、-、\*、/、num

非终结符：E          开始符号：E

**\*一个 1 分**

（2）不是，因为前缀表达式无需括号即可表达正确计算顺序

**\*结论 1 分，解释 2 分**

（3）

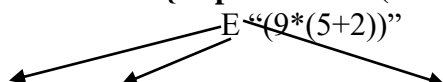
$E \rightarrow + E_1 E_2 \quad \{ E.\text{prefix} = "(" \parallel E_1.\text{prefix} \parallel "+" \parallel E_2.\text{prefix} \parallel ")"; \}$

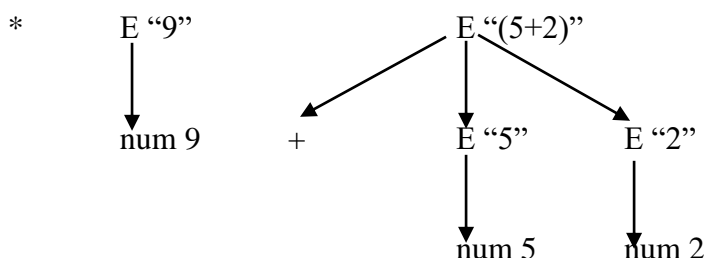
$E \rightarrow - E_1 E_2 \quad \{ E.\text{prefix} = "(" \parallel E_1.\text{prefix} \parallel "-" \parallel E_2.\text{prefix} \parallel ")"; \}$

$E \rightarrow * E_1 E_2 \quad \{ E.\text{prefix} = "(" \parallel E_1.\text{prefix} \parallel "*" \parallel E_2.\text{prefix} \parallel ")"; \}$

$E \rightarrow / E_1 E_2 \quad \{ E.\text{prefix} = "(" \parallel E_1.\text{prefix} \parallel "/" \parallel E_2.\text{prefix} \parallel ")"; \}$

$E \rightarrow \text{num} \quad \{ E.\text{prefix} = \text{itos}(\text{num.val}); \}$





\*语法制导定义 5 分，注释语法树 5 分

若设计的是翻译模式（打印数、运算符和括号，而不是计算中缀字符串保存在属性中），完全正确（翻译模式、带语义动作的语法树以及最后的中缀表达式都正确给出了）给 7 分。

部分正确的情况在上述基础上酌情扣分。

得 分

五、（14 分）对下面已拓广的文法：

(1) 基于 LR(1)项目给出能识别活前缀集合的 NFA（7 分）

(2) 用此 NFA 实现对句子 cdccd 的语法分析（仍采用(状态栈，输入缓冲区)的描述方式）（7 分）。

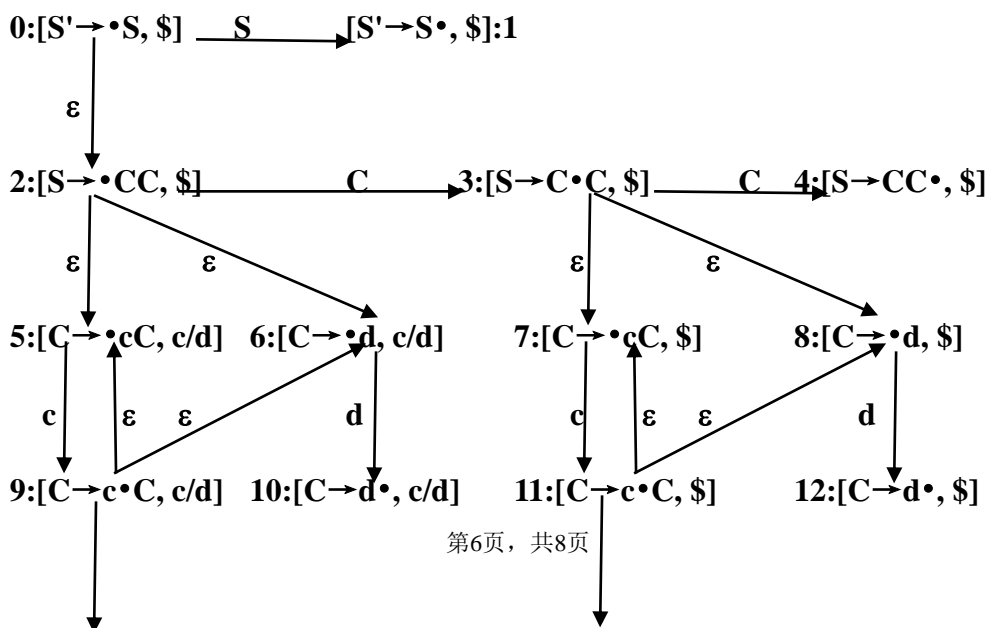
$S \rightarrow S'$

$S \rightarrow CC$

$C \rightarrow cC \mid d$

答：

(1) NFA 如下



C

C

13:[ $C \rightarrow cC \bullet$ ,  $c/d$ ]14:[ $C \rightarrow cC \bullet$ , \$]

\*计算 LR(1)项目集规范族，正确给出了 DFA 扣 1 分，计算 LR(0)项目集规范族，正确给出了 DFA 扣 3 分。  
部分正确的情况在上述基础上酌情扣分。

(2) 分析过程如下：

状态栈	输入缓冲区	归约
$\{0,2,5,6\}$	cdccd\$	
$\{0,2,5,6\}c\{5,6,9\}$	dccd\$	
$\{0,2,5,6\}c\{5,6,9\}d\{10\}$	ccd\$	$C \rightarrow d$
$\{0,2,5,6\}c\{5,6,9\}C\{13\}$	ccd\$	$C \rightarrow cC$
$\{0,2,5,6\}C\{3,7,8\}$	ccd\$	
$\{0,2,5,6\}C\{3,7,8\}c\{7,8,11\}$	cd\$	
$\{0,2,5,6\}C\{3,7,8\}c\{7,8,11\}c\{7,8,11\}$	d\$	
$\{0,2,5,6\}C\{3,7,8\}c\{7,8,11\}c\{7,8,11\}d\{12\}$	\$	$C \rightarrow d$
$\{0,2,5,6\}C\{3,7,8\}c\{7,8,11\}c\{7,8,11\}C\{14\}$	\$	$C \rightarrow cC$
$\{0,2,5,6\}C\{3,7,8\}c\{7,8,11\}C\{14\}$	\$	$C \rightarrow cC$
$\{0,2,5,6\}C\{3,7,8\}C\{4\}$	\$	$S \rightarrow CC$
$\{0,2,5,6\}S\{1\}$	\$	accept

\*上一问计算 LR(1)项目集规范族，且分析过程正确的扣 1 分，计算 LR(0)项目集规范族，且分析过程正确的扣 3 分。  
部分正确的情况在上述基础上酌情扣分。

