

《编译技术》期中考试试卷（2017 年春季）

考试时间：2017 年 4 月 19 日

姓名：\_\_\_\_\_ 学号：\_\_\_\_\_

题号	一	二	三	四	五	六	总分
分数							
阅卷人							

**考 场 纪 律**

1. 请持学生证入场考试，并按指定座位就座；除必要的文具和教师指定的用具用书外，其他所有物品包括手机、呼机、MP3、电子词典、书籍、笔记、纸张等严禁带入座位，必须放在指定位置。凡有试题印制问题请向监考教师提出，不得向其他考生询问。
2. 认真、诚实、独立并在规定时间内完成答卷，严禁任何形式的违纪作弊行为；否则，本答卷成绩以 0 分记，并根据《北京大学本科考试工作与学术规范条例》给予纪律处分。
3. 提前交卷的考生不要在考场逗留，不要在门口、窗外大声喧哗。考试结束时间到，请停止答卷，在座位等候监考教师收卷并清点完毕，方可离开考场；考题和试卷不得带出考场。

以下为试题和答题纸，共 13 页。

一、单选题（20 分，每小题 2 分）。请把答案填入下面表格中。

1	2	3	4	5	6	7	8	9	10

- 文法产生的全部\_\_是该文法所描述的语言：  
(A) 句子 (B) 终结符号 (C) 句型 (D) 非终结符号
- 对 LR 分析表的构造，不可能存在以下何种动作冲突？  
(A) 移进/归约  
(B) 移进/移进  
(C) 归约/归约  
(D) 以上都不对
- 正则表达式  $(a|b|c|\varepsilon)(a|b|c|\varepsilon)(a|b|c|\varepsilon)(x|y|\varepsilon)$  表示的语言中包含的互不相同的**非空串**有多少个：  
(A) 54 (B) 84 (C) 119 (D) 192
- 给定文法： $E \rightarrow E * E | E + E | (E) | \text{int}$   
对于字符串  $5 * 5 + (3 * 2) + 1$  可以推导出多少种分析树？  
(A) 1 (B) 4 (C) 5 (D) 8
- 下列文法中，有二义性的是：  
(A)  $S \rightarrow S[S] | S * S | S == S | (S) | \text{id}$   
(B)  $S \rightarrow S : S | \text{id}$   
(C)  $S \rightarrow L | S : L ?$   
 $L \rightarrow \text{id} | (S)$   
(D)  $S \rightarrow 0S | 0S1S | \varepsilon$
- 给定文法： $S \rightarrow SS+ | SS* | a$ ，则右句型  $SSa^*+aa^*$  的句柄是：  
(A) S (B)  $SS+$  (C)  $SS^*$  (D) a
- 现有句型  $\gamma a \beta b w$ 。其最右推导是  $S \Rightarrow \dots \Rightarrow \gamma A b w \Rightarrow \gamma a \beta b w$ ，最后一步用的产生式是  $A \rightarrow a \beta$ 。试分别指出 LR(1)方法和 LL(1)方法在扫描到句型  $\gamma a \beta b w$  的什么位置时决定用产生式  $A \rightarrow a \beta$ ？

- (A)  $\beta$  和 b                      (B) w 和  $\beta$   
 (C) a 和 b                        (D) b 和 a

8. 下面的文法是一个算符优先的文法：

$E \rightarrow E+T \mid T$

$T \rightarrow a * T \mid a$

使用移进-归约分析，接受输入串  $a+a*a$  分别需要移进和归约各多少步？

- (A) shift=5, reduce=5  
 (B) shift=5, reduce=8  
 (C) shift=6, reduce=5  
 (D) shift=6, reduce=8

9. 设有文法  $G[S]$ ：

$S \rightarrow aTUb \mid \epsilon$

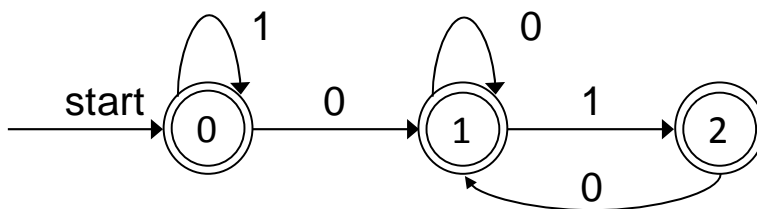
$T \rightarrow cUc \mid bUb \mid aUa$

$U \rightarrow Sb \mid cc$

下面关于这个文法的 First 集合和 Follow 集合的说法正确的是：

- (A) S 的 First 集是  $\{\epsilon, a, b\}$   
 (B) U 的 First 集是  $\{a, c\}$   
 (C) T 的 Follow 集是  $\{a, b, c\}$   
 (D) S 的 Follow 集是  $\{\$, b, c\}$

10. 下面的正则表达式中哪个可以表示图中的 DFA？



- (A)  $1^*(0 \mid 10)^*(0 \mid 1 \mid \epsilon)$   
 (B)  $1^*(10 \mid 01)^*(0 \mid 1 \mid \epsilon)$   
 (C)  $1^*0(0 \mid 10)^*$   
 (D)  $1^*0(0 \mid 01)^*$

二、多选题（20 分，每小题 2 分）。请把答案填入下面表格中。

1	2	3	4	5	6	7	8	9	10

1. 下列说法不正确的是：

- (A) 对于同一个文法，SLR 和 LR(0)的项集族是略有不同的
- (B) 自底向上的语法分析，必须要先消除左递归
- (C) 句柄指的是最左推导出每个句型所用的最后一个产生式的产生式体
- (D) LALR 比 SLR 能处理更多的文法，但需要构造的语法分析表却不会更大

2. 考虑如下文法：

$S \rightarrow A * B$

$A \rightarrow B \mid C$

$B \rightarrow (C)$

$C \rightarrow B + C \mid D$

$D \rightarrow 1 \mid 0$

在该文法的基础上添加下面哪一个文法会产生左递归？

- (A)  $A \rightarrow D$
- (B)  $D \rightarrow S$
- (C)  $C \rightarrow A + C$
- (D)  $C \rightarrow 1C$

3. 给定语言 L：字母表  $\Sigma = \{a, b\}$  上所有不含子串 bba 的 a 和 b 的串。则不等价于该语言的正则表达式是？

- (A)  $(b|ba)^*b^*$
- (B)  $(a|ba)^*b^*$
- (C)  $(a|bab)^*b^*$
- (D)  $(b|ab)^*b^*$

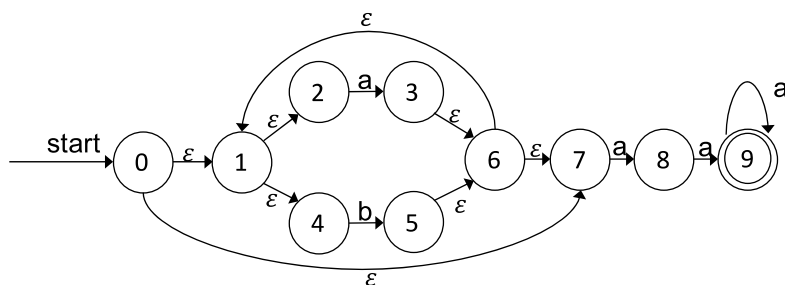
4. 以下说法正确的是：

- (A) LL(k)表示向前预测 k 步的一种自顶向下的分析方法。
- (B) LL(k)每次向前看 k 个输入符号的一种自顶向下分析方法。
- (C) LR(k)表示从右到左分析，每次决定动作时向前看 k 个输入符号的一种自底向上分析方法。

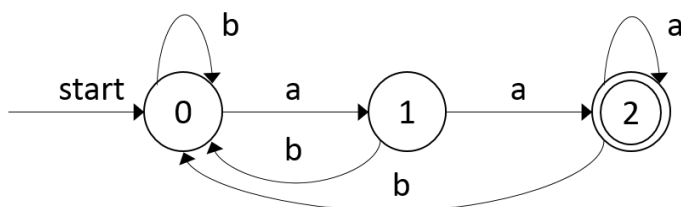
(D) LR(k)表示从左到右分析，每次决定动作时向前看 k 个输入符号的一种自底向上分析方法。

5. 等价于正则表达式  $(a|b)^*aa$  的 DFA 或 NFA 是：

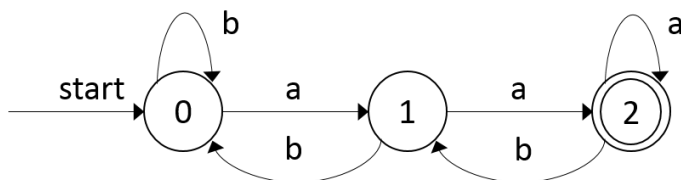
(A)



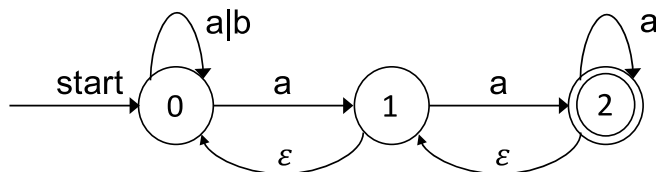
(B)



(C)



(D)



6. 考虑如下 LL(1)文法:

$S \rightarrow T d (A)$

$T \rightarrow d F$

$F \rightarrow *F \mid \varepsilon$

$A \rightarrow L \mid \varepsilon$

$L \rightarrow T d M$

$M \rightarrow ,L \mid \varepsilon$

则对于 LL(1)预测分析表, 以下说法正确的是:

(A) 在表中, 输入符号为'd'时 T 对应的表项为  $T \rightarrow d F$

(B) 在表中, 输入符号为 '(' 时 A 对应的表项为  $A \rightarrow L$

(C) 在表中, 输入符号为 ')' 时 M 对应的表项为  $M \rightarrow \varepsilon$

(D) 在表中, 输入符号为'd'时 F 对应的表项为  $F \rightarrow \varepsilon$

7. 以下哪个文法所产生的语言可以同样被正则表达式表达?

(A)  $A \rightarrow aA \mid b$

(B)  $A \rightarrow aaAb \mid \varepsilon$

(C)  $A \rightarrow (B) \mid (BB)$

$B \rightarrow (CC) \mid (CCC)$

$C \rightarrow (DDD)$

$D \rightarrow ()$

(D)  $A \rightarrow AAaab \mid \varepsilon$

8. 下面哪些文法是非二义性的, 并且和以下文法:

$E \rightarrow E+E \mid E-E \mid E*E \mid E/E \mid d$

等价?

(A)  $E \rightarrow d+E \mid d-E \mid d*E \mid d/E \mid d$

(B)  $E \rightarrow d+E' \mid d-E' \mid d+E \mid d-E$

$E' \rightarrow d*E' \mid d/E' \mid d$

(C)  $E \rightarrow E+d \mid E-d \mid E*d \mid E/d \mid d$

(D)  $E \rightarrow E'+E \mid E'-E \mid E'$

$E' \rightarrow d*E' \mid d/E' \mid d$

9. 考虑如下扩展文法:

$S' \rightarrow S$

$S \rightarrow X$

$X \rightarrow DE$

$D \rightarrow PD \mid \varepsilon$

$E \rightarrow QE \mid \varepsilon$

$P \rightarrow id\ id$

$Q \rightarrow id = id$

下面哪些条目在该文法的 LR(0)分析状态机的初始状态集合中?

(A)  $S \rightarrow \bullet X$

(B)  $D \rightarrow \bullet$

(C)  $E \rightarrow \bullet QE$

(D)  $Q \rightarrow \bullet id = id$

10. 在上一题目 (题目 9) 文法的基础上, 下面说法不正确的是:

(A) 该文法的 LR(0)分析状态机有 14 个状态

(B) 该文法属于 SLR(1)

(C) 该文法属于 LALR(1)

(D) 该文法属于 LR(1)

### 三、简答题（10 分）。

1. 该正则表达式  $b^*(ab^*ab^*)^*$  表示的是什么语言？简单说明原因。（3 分）

2. 考虑如下文法：

$S \rightarrow T | S+T$

$T \rightarrow U | T/U$

$U \rightarrow (S) | a$

(1) 给出句子  $a + (a / a) + a$  的最左推导并给出相应的语法树。（3 分）

(2) 写出句型  $S + T / (T + a / a)$  的最右推导，并指出该句型的句柄。（4 分）



四、正则表达式与自动机（15 分，每小题 5 分）

现有能某种语言的正则表达式  $(a \mid \varepsilon)(a \mid b)^*(aba)^*bb$

(1) 请构造出对应的 NFA。

(2) 请将 NFA 转化为 DFA。

(3) 请将 DFA 化简为最简 DFA。

五、LL 分析（15 分，每小题 5 分）

(1) 将文法  $G[S]$  改写为等价的  $G'[S]$ ，使  $G'[S]$  不含左递归和左公共因子。

$G[S]: S \rightarrow A$   
 $A \rightarrow AB-- / AB++ / idB$   
 $B \rightarrow B-- / B++ / .$

(2) 为  $G'[S]$  中的非终结符号计算相应的 First 和 Follow 集合。  
(行 / 列数不够时请自行添加)

	First	Follow
S		

(3) 为  $G'[S]$  构造相应的 LL 分析表。(行 / 列数不够时请自行添加)

	id	.	-	+	\$
S					

## 六、LR 分析（20 分，每小题 4 分）

对第五题中的  $G[S]$  进行拓广后，得到  $G'[s]$ ：

- (0)  $S' \rightarrow S$
- (1)  $S \rightarrow A$
- (2)  $A \rightarrow AB--$
- (3)  $A \rightarrow AB++$
- (4)  $A \rightarrow idB$
- (5)  $B \rightarrow B--$
- (6)  $B \rightarrow B++$
- (7)  $B \rightarrow .$

(1) 构造该文法对应的 LR(0)项集族和相应的 DFA。

(2) 构造上述文法对应的 LR(1)项集族。如果存在同心项，指出是哪些项集。



(5) 用上题的得到的语法分析表对字符串 `id . . + + - - - -` 进行分析，填写直到字符串被接受或报错时，栈的状态、采取的动作等。  
(行 / 列数不够时请自行添加)

栈[最左为栈底]	符号	输入	动作