

《编译技术》期中考试试卷（2018 年春季）

考试时间：2018 年 4 月 18 日

姓名：_____ 学号：_____

题号	一	二	三	四	五	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、单项选择题（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. 以下说法正确的是：

- (A) NFA 的表达能力比 DFA 更强，因为 DFA 只是 NFA 的一个特例
- (B) 文法 $S \rightarrow aS \mid bS \mid ab$ 接受由 a、b 组成的任意长度的字符串
- (C) 语法分析必须先消除左递归
- (D)** 上下文无关文法能描述所有正则语言

2. 如果文法 G 是无二义的，则它的任何句子 α _____。

- (A) 可能没有最左推导或最右推导
- (B)** 最左推导和最右推导对应的语法树必定相同
- (C) 可能存在两个不同的最左推导，但它们对应的语法树相同
- (D) 最左推导和最右推导必定相同

3. 对 LR(0), LR(1), LALR(1), SLR(1) 所能处理的文法的集合之间关系，表述正确的是

- (A) $SLR(1) \supset LR(1) \supset LALR(1) \supset LR(0)$
- (B) $LR(1) \supset LALR(1) \supset SLR(1) \supset LR(0)$
- (C)** $LR(1) \supset SLR(1) \supset LALR(1) \supset LR(0)$
- (D) $LR(0) \supset LR(1) \supset LALR(1) \supset SLR(1)$

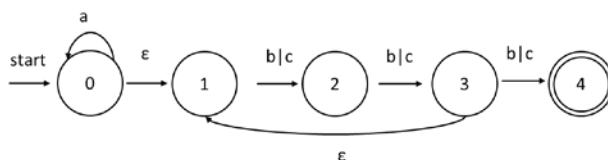
82 4. 正则表达式 $(a \mid b \mid c \mid d \mid \epsilon)(a \mid b \mid c \mid \epsilon)(a \mid b \mid \epsilon)(a \mid b)$ 表示的语言包含多少个互不相同的字符串？

- (A) 48
- (B) 89
- (C) 71
- (D) 83

5. 串 “abcdefg” 的非空子串有多少个？

- (A) 127
- (B) 128
- (C) 27
- (D)** 28

6. 与下图中 NFA 等价的正则表达式是：



- (A) $(a(b|c)(b|c))^*(b|c)$
- (B) $a^*(b|c)^+$
- (C) $a^*(b|c)^+(b|c)$
- (D) $a^*(bc|cb|cc|bb)^+(c|b)$

7. 文法 $E \rightarrow E+E \mid E^*E \mid \langle E, E \rangle \mid i$ 的句子 $i^* \langle i+i^*i, i^*i+i \rangle + i^*i$ 有多少棵不同的语法树？

- (A) 13
- (B) 24
- (C) 20
- (D) 18

8. 考虑如下的文法：

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

在该文法的基础上添加下面哪一个产生式会产生左递归？

- (A) $D \rightarrow A$
- (B) $D \rightarrow B$
- (C) $B \rightarrow D$
- (D) $A \rightarrow D$

9. 设有文法 $G[S]$:

$S \rightarrow \{SA\} \mid a$
 $A \rightarrow +SB \mid Bb \mid \epsilon$
 $B \rightarrow -SAc \mid \epsilon$

则 S 的 Follow 集合是：

- (A) $\$, \}, +, -, a, b, c$
- (B) $\$, \}, +, -$
- (C) $\$, \}, +, -, b, c$
- (D) $\$, \}, +, -, b$

10. 对于文法

$A \rightarrow AxBy \mid yA \mid z$
 $B \rightarrow xA \mid y$

下列哪个文法生成的语言和该文法一样？

- (A) $A \rightarrow yAD \mid zD$
 $B \rightarrow xAyD \mid yD$
 $D \rightarrow xB \mid \varepsilon$
- (B) $A \rightarrow yAD \mid zD$
 $B \rightarrow xAyD \mid yyD$
 $D \rightarrow xB \mid \varepsilon$
- (C) $A \rightarrow yAD \mid zD$
 $B \rightarrow xAy \mid yD$
 $D \rightarrow xB \mid \varepsilon$
- (D) $A \rightarrow yAD \mid zD$
 $B \rightarrow xAy \mid yyD$
 $D \rightarrow xB \mid \varepsilon$

二、不定项选择题（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

- 下列说法正确的是：
 - 文法 G 的一个句子存在多个推导，则 G 是二义性的
 - 因为 $SLR(1)$ 可以消除 $LR(0)$ 中所有的移进-归约冲突，所以它可以比 $LR(0)$ 处理更多文法
 - 串 S 的任意一个子串必定也是串 S 的子序列
 - LR 中的 R 表示自右向左读入输入串
- 以下说法不正确的是：
 - 正则表达式和上下文无关文法表达能力相同
 - 正则表达式和 NFA 表达能力相同
 - 解释型编译器和编译型编译器的区别在于，解释型编译器并没有生成可执行的机器码
 - 任意文法中，产生式规则的左部都是非终结符号
- 设有文法

$$S \rightarrow A \mid B$$

$$A \rightarrow x \mid xA \mid xC \mid yAA$$

$$B \rightarrow y \mid yB \mid yC \mid xBB$$

$$C \rightarrow xCy \mid yCx \mid CC \mid \varepsilon$$
 可以推导出的串有？
 - xyxyxyy
 - yyyxxx
 - yxy
 - yyxyxxx
- 给定文法 $S \rightarrow (S, S) \mid aSc \mid SS\& \mid b$ ，则句柄是 $SS\&$ 的右句型有：
 - $(S, aSb\&c)$
 - $S(S, aSSS\&\&c)\&$
 - $a(aSS\&c, aSS\&c)$
 - $S(S, SS\&)\&$

5. 考虑如下文法:

$E \rightarrow E + T \mid T$

$T \rightarrow T \times F \mid F$

$F \rightarrow (E) \mid id$

对于字符串 $id \times id$, 下面哪些是推导该字符串过程中产生的右句型?

(A) $F \times id$

(B) $F \times F$

(C) $T \times F$

(D) $id \times F$

6. 设有以下文法 $G(S)$

$S \rightarrow 2AB$

$A \rightarrow 0A \mid 0$

$B \rightarrow B1 \mid \varepsilon$

下列说法正确的有:

(A) 该文法包含左递归

(B) 该文法描述的语言为 $L(G) = \{20^n 1^n \mid n \geq 0\}$

(C) 该文法不包含左公因子

(D) 该文法能够用正则表达式表达

7. 考虑如下文法:

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

$A \rightarrow S \mid SB \mid x \mid \varepsilon$

$B \rightarrow SB \mid y$

下面哪些项在该文法的 SLR 分析状态机的初始状态集合中?

(A) $A \rightarrow \bullet SB$

(B) $S \rightarrow \bullet$

(C) $B \rightarrow \bullet SB$

(D) $A \rightarrow \bullet x$

8. 在上一题目 SLR 状态机的初始状态集合的基础上, 下面说法正确的有:

(A) 在该状态下对于输入符号 x 有移进-规约冲突

(B) 在该状态下对于输入符号 $($ 有移进-规约冲突

(C) 在该状态下对于输入符号 $($ 有规约-规约冲突

(D) 在该状态下对于转换 S 有规约-规约冲突

9. 下列哪些文法表示同一种语言?

(A) $S \rightarrow A \mid B$

$A \rightarrow 1 \mid 1A \mid 1E \mid 0AA$

- $$\begin{aligned}
 & B \rightarrow 0 \mid 0B \mid 0E \mid 1BB \\
 & E \rightarrow 0E1 \mid 1E0 \mid EE \mid \varepsilon \\
 (B) \quad & S \rightarrow A \mid B \\
 & A \rightarrow 1 \mid 0A1 \mid 1A0 \mid AA \mid EAE \\
 & B \rightarrow 0 \mid 0B1 \mid 1B0 \mid BB \mid EBE \\
 & E \rightarrow 0E1 \mid 1E0 \mid EE \mid \varepsilon \\
 (C) \quad & S \rightarrow TS \mid C \mid D \\
 & T \rightarrow 1B \mid 0A \\
 & A \rightarrow 1 \mid 0AA \\
 & B \rightarrow 0 \mid 1BB \\
 & C \rightarrow 0E \\
 & E \rightarrow 0E \mid 0E1E \mid \varepsilon \\
 (D) \quad & S \rightarrow AB \\
 & A \rightarrow 1 \mid 1A \mid 1E \mid 0AA \mid \varepsilon \\
 & B \rightarrow 0 \mid 0B \mid 0E \mid 1BB \mid \varepsilon \\
 & E \rightarrow 0E1 \mid 1E0 \mid EE \mid \varepsilon
 \end{aligned}$$

10. 考虑如下扩展文法:

$$S' \rightarrow S$$

$$S \rightarrow CC$$

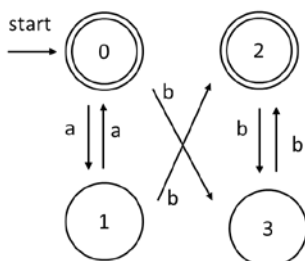
$$C \rightarrow cC \mid d$$

对输入字符串“ccdc d”做 LALR 分析, 当用 $C \rightarrow cC$ 规约时, 栈中可能有几个字符 (不包含状态)?

- (A) 1 (B) 2 (C) 3 (D) 4

三、简答题（25 分）。

1. 抽象的描述下面的 DFA 接受的语言，并给出一个等价的正则表达式。（4 分）



2. 设有如下两个文法：

G1:

$S \rightarrow S - S \mid A$

$A \rightarrow A \wedge B \mid B$

$B \rightarrow a$

G2:

$S \rightarrow S - S$

$S \rightarrow S \wedge S$

$S \rightarrow a$

- (1) 分别根据这两个文法写出 $a \wedge a - a$ 的最右推导，并指出倒数第二步推导用到的句柄。（2 分）

- (2) 这两个文法中，哪些适合 LL(1)分析？（2 分）

3. 设有正则表达式 $(1(0^*1^*)^*1^+)|0$, 请构造出其 mini (最简) DFA。
(5 分)
4. 构造一个上下文无关文法, 使其接收如下语言 $L = \{a^n b^m c^k : n = m \text{ 或 } m \leq k\}$ 。请解释思路。(5 分)
5. 为语言 $L = \{u\omega\omega^R v : u, v, \omega \in \{a, b\}^+, \omega^R \text{ 是 } \omega \text{ 的回文串}\}$ 构造一个正则表达式。(7 分)

四、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。(digit 视为一个终结字符)

$G[S]$: $S \rightarrow (S) | S * A | A$
 $A \rightarrow A + B | B$
 $B \rightarrow digit B | digit | \epsilon$

(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。 $G'[S]$ 是 LL(1) 文法吗? 请作出说明。(行 / 列数不够时请自行添加)

	digit	*	+	()	\$
S						

五、LR 分析（20 分，每小题 4 分）

现有如下增广文法 $G'[s]$:

- (0) $S' \rightarrow S$
- (1) $S \rightarrow DE$
- (2) $D \rightarrow PD$
- (3) $D \rightarrow \varepsilon$
- (4) $E \rightarrow QE$
- (5) $E \rightarrow \varepsilon$
- (6) $P \rightarrow xx$
- (7) $Q \rightarrow y=y$

(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。

(3) 为该文法构造 LR(1)分析表。(行 / 列数不够时请自行添加)

状态	ACTION				GOTO				
	x	y	=	\$	S	D	E	P	Q
0									
1									
2									
3									
4									

(4) 该文法是 SLR 的么？是 LALR 的么？请详细说明判断理由。

(5) 用第(3)小题得到的语法分析表对字符串 $xxxxy=yy=y$ 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)

栈中状态	栈中符号	输入	动作
0		$xxxxy=yy=y\$$	