

《SE-303 编译原理》期末试题 (A 卷答案卷)

(考试形式：闭卷 考试时间：2 小时)

Part one: Answer the following questions (15 points. 5 points for each item.)

1. What phases does a compiler have?

【参考答案】

编译程序分为：词法分析、语法分析、语义分析、中间代码生成、代码优化、代码生成、符号表管理和出错处理八个阶段。

【评分标准】

本小题 5 分。

词法分析、语法分析、语义分析、中间代码生成、代码优化、代码生成各 1 分；

由于符号表管理和出错处理可以看成是多个阶段共有的部分，因此如果没有答出这两部分建议也不扣分；

如果按照课本将代码优化分成机器无关的优化和机器相关的优化，也可以给分。

2. What are the two main techniques in syntax analysis?

【参考答案】

语法分析的两个主要技术是自顶向下分析技术和自底向上分析技术。

【评分标准】

本小题 5 分。若具体列举分析技术，如 LL(1), LR(0), SLR(1) 等，没有概括，则给 4 分。

3. Describe the four components of a context-free grammar.

【参考答案】

上下文无关文法由终结符号、非终结符号、一个开始符号与一组产生式组成。

【评分标准】

本小题 5 分。

终结符号、非终结符号、一个开始符号各 1 分；如果未明确说明是“一个”，扣 0.5 分。一组产生式占 2 分。

Part two: Compute and answer the following questions (85 points)

4. (15 points) Convert the following NFA into a DFA with minimum number of states:

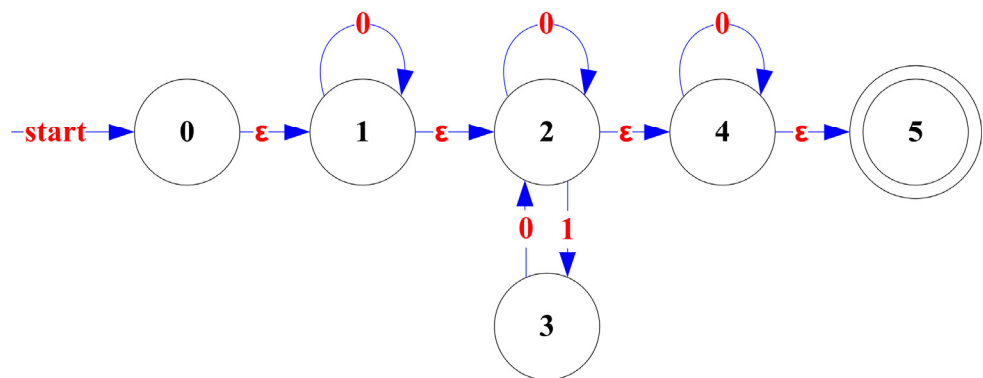


图 4-1

【参考答案】

对图 4-1 的 NFA 先转换为 DFA 如下图 4-2 所示。

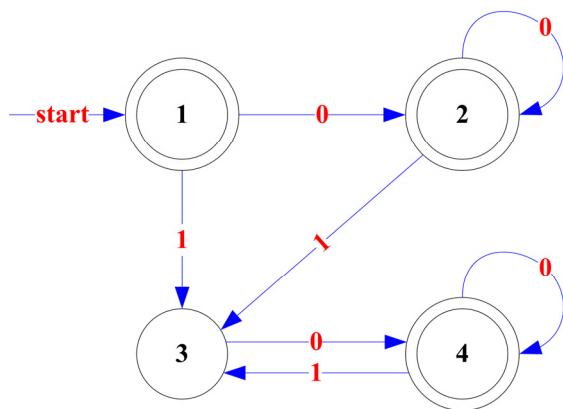


图 4-2

具体的转换过程如表 4-1 所示

表 4-1

NFA	DFA	0	1
{0,1,2,4,5}	1	2	3
{1,2,4,5}	2	2	3
{3}	3	4	Φ
{2,4,5}	4	4	3

【评分标准】

本小题 7 分。

如图 4-2 所示：转换后的 DFA 正确，但状态数多于 4 或转换关系不完全正确，得 5 分。

如果答案不正确，但能写出如表 4-1 所示的求解过程，可酌情给分。

(2)

【参考答案】

将图 4-2 的 DFA 进行最小化优化，优化后的 DFA 如图 4-3 所示，转换状态如表 4-2 所示。

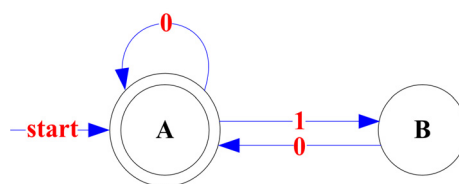


图 4-3

表 4-2

DFA	优化后 DFA 状态
1,2,4	A
3	B

【评分标准】

本小题 8 分。

如图 4-3 所示：优化后的 DFA 正确，但状态数多于 2 或转换关系不完全正确，得 5 分；

DFA 正确，状态数等于 2 且状态转换关系正确，得 8 分。

如果按以上转换过程进行分析（如图 4-3 所示，表 4-2 所示），但最终结果不对，可根据计算过程酌情给分。

5. (10 points) Which of the following language is a regular language, a context-free language or a context-sensitive language? For the regular language, write its regular expression, for the context-free language (which is not a regular language), write its context-free grammar.

(1) $L_1 = \{w \mid w \in \{0,1\}^* \text{ and } w \text{ does not contain subsequence } 011\}$;

(2) $L_2 = \{wcw \mid w \in \{0,1\}^* \text{ and } c \neq 0, 1\}$;

(3) $L_3 = \{w \mid w \in \{0,1\}^* \text{ and } w = 1^n 0^m 1^m 0^n, n \geq 0, m \geq 0\}$.

(1)

【参考答案】是正则语言。正则表达式为： $1^*0^*(1|\epsilon)0^*$

【评分标准】本小题 4 分。答对正则语言得 2 分，给出正确的正则表达式得 2 分。

(2)

【参考答案】是上下文有关语言。

【评分标准】本小题 2 分。答对上下文有关语言得 2 分。（希望同学能说明此题是上下文有关语言的原因）

(3)

【参考答案】是上下文无关语言，上下文无关方法为：

$S \rightarrow 1S0|A$

$A \rightarrow 0A1|\epsilon$

【评分标准】本小题 4 分。答对上下文无关语言得 2 分，给出正确的文法得 2 分。

6. (10 points) Given the following grammar:

G: $S \rightarrow SS|(S)|()$

(1) Prove this grammar is ambiguous.

- (2) Describe what language this grammar generate.
- (3) Construct an unambiguous grammar that generate the same language as this grammar.

(1)

【参考答案】

对于句子000, 可以构造出两棵语法分析树, 如图 6-1:

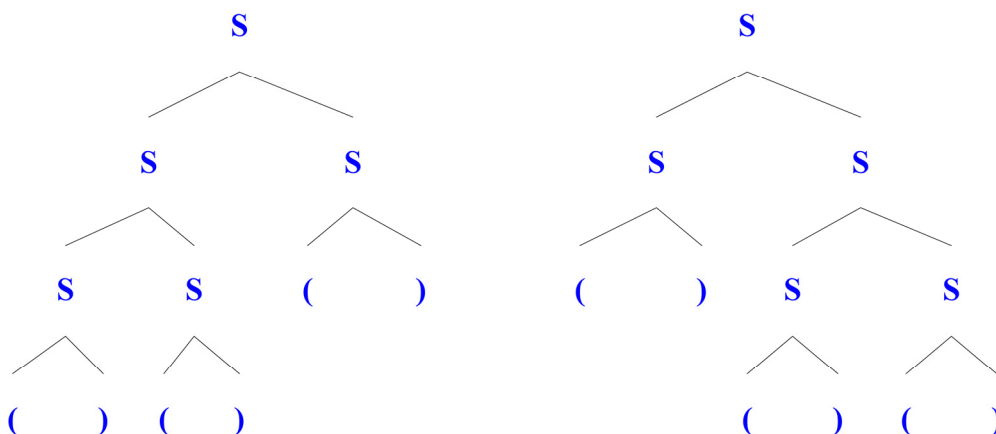


图 6-1

(2)

【参考答案】

产生所有由 '(' 和 ')' 组成的匹配括号串.

【评分标准】本小题 3 分.

【评分标准】本小题 4 分.正确给出证明过程即可得全分。

(3)

【参考答案】

$S \rightarrow ST \mid T$

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

【评分标准】本小题 3 分.正确给出无二义性文法即可得全分。给出文法与以上文法不同，酌情给分。

7. (10 points) Construct the predictive parsing table of the following grammar.(Hint: To eliminate left recursion of the grammar first)

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

$T \rightarrow T \bullet F \mid F$

$F \rightarrow F* \mid a \mid b$

【参考答案】

(1) 消除左递归，得文法

$E \rightarrow TE'$

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

$T \rightarrow FT'$

$T' \rightarrow \bullet FT' \mid \epsilon$

$$F \rightarrow aF' | bF'$$

$$F' \rightarrow *F' \mid \varepsilon$$

【评分标准】

答对此部分可得 4 分。

由产生式知：

$$\begin{aligned} \text{FIRST}(E) &= \{a, b\} & \text{FOLLOW}(E) &= \{\$ \} \\ \text{FIRST}(E') &= \{+, \varepsilon\} & \text{FOLLOW}(E') &= \{\$ \} \\ \text{FIRST}(T) &= \{a, b\} & \text{FOLLOW}(T) &= \{+, \$\} \\ \text{FIRST}(T') &= \{\bullet, \varepsilon\} & \text{FOLLOW}(T') &= \{+, \$\} \\ \text{FIRST}(F) &= \{a, b\} & \text{FOLLOW}(F) &= \{\bullet, +, \$\} \\ \text{FIRST}(F') &= \{*, \varepsilon\} & \text{FOLLOW}(F') &= \{\bullet, +, \$\} \end{aligned}$$

【评分标准】

答对此部分可得 5 分。

预测分析表如下表 7-1 所示：

表 7-1

	+	•	*	a	b	\$
E				$E \rightarrow TE'$	$E \rightarrow TE'$	
E'	$E' \rightarrow +TE'$					$E' \rightarrow \varepsilon$
T				$T \rightarrow FT'$	$T \rightarrow FT'$	
T'	$T' \rightarrow \varepsilon$	$T' \rightarrow \bullet FT'$				$T' \rightarrow \varepsilon$
F				$F \rightarrow aF'$	$F \rightarrow bF'$	
F'	$F' \rightarrow \varepsilon$	$F' \rightarrow \varepsilon$	$F' \rightarrow *F'$			$F' \rightarrow \varepsilon$

【评分标准】

答对此部分可得 6 分。

8. (15 points) Given the following grammar:

$$A \rightarrow aAd$$

$$A \rightarrow aAb$$

$$A \rightarrow \varepsilon$$

(1) Is the grammar an SLR(1) grammar? Explain the reason concisely.

(2) Given the input string: ab#, please describe the parsing process in detail.

(1)

【参考答案】

对于文法： $A \rightarrow aAd|aAb|\varepsilon$ ，其拓广文法为 G' ，增加产生式 $S' \rightarrow A$ ，设产生式排序为：

$$\begin{aligned} (0) S' &\rightarrow A \\ (1) A &\rightarrow aAd \\ (2) A &\rightarrow aAb \\ (3) A &\rightarrow \varepsilon \end{aligned}$$

【评分标准】

答对此部分可得 2 分。

由产生式知：

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

$\text{First}(A) = \{\epsilon, a\}$

$\text{Follow}(S') = \{\$\}$

$\text{Follow}(A) = \{d, b, \$\}$

【评分标准】

答对此部分可得 2 分。

G' 的 LR(0) 项目集族及识别活前缀的 DFA 如下图 7-1 所示：

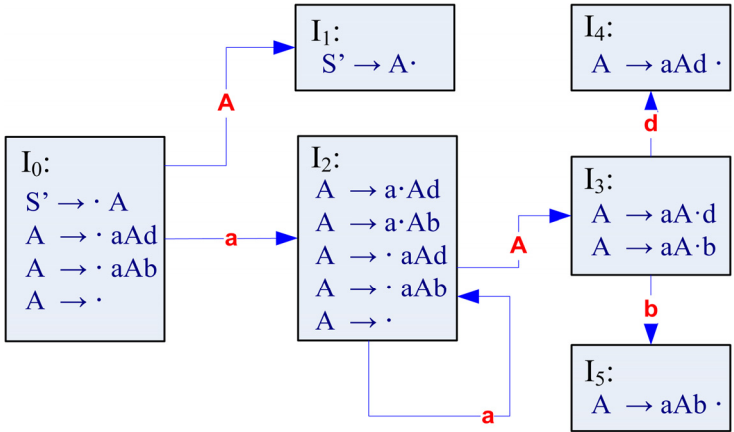


图 8-1

在 I_0 中： $A \rightarrow \cdot aAd$ 和 $A \rightarrow \cdot aAb$ 为移进项目， $A \rightarrow \cdot$ 为归约项目，存在移进-归约冲突，因此所给文法不是 LR(0) 文法。

在 I_0 与 I_2 中： $\text{Follow}(A) \cap \{a\} = \{d, b, \$\} \cap \{a\} = \emptyset$ ，所以在 I_0 、 I_2 中的移进-归约冲突可以由 Follow 集解决，所以 G 是 SLR(1) 文法。

【评分标准】

答对此部分可得 6 分。如果答案不正确，但能正确地给出分析过程，可酌情给分。

(2)

【参考答案】

构造的 SLR(1) 分析表如下表 8-1 所示：

表 8-1

State	Action				Goto
	a	d	b	\$	A
0	S2	r3	r3	r3	1
1				Acc	
2	S2	r3	r3	r3	3
3		S4	S5		
4		r1	r1	r1	
5		r2	r2	r2	

对输入串 $ab\#$ 的分析过程如下表 7-2 所示：

表 8-2

状态栈	文法符号栈	剩余输入串	动作
0	#	ab#	移入
0 2	#a	b#	规约: $A \rightarrow \epsilon$
0 2 3	#aA	b#	移入
0 2 3 5	#aAb	#	规约: $A \rightarrow aAb$
0 1	#A	#	

【评分标准】

答对此部分可得 5 分。如果答案不正确，但能正确地给出分析过程，可酌情给分。

9. (Optional 1) (15 points)

Based on the syntax-directed definitions in the textbook, translate the following statement into quadruple (three-address statement) sequence:

```
WHILE A < C  ∧  A < D DO
  IF A = 1 THEN
    A := A+1
  ELSE WHILE A < D DO A := A+2
```

【参考答案】

- (1) (j<, A, C, 3)
- (2) (j, _, _, 0)
- (3) (j<, A, D, 5)
- (4) (j, _, _, 2) ← S.CHAIN
- (5) (j=, A, 1, 7)
- (6) (j, _, _, 10)
- (7) (+, A, 1, T_1)
- (8) (:=, T_1 , _, A)
- (9) (j, _, _, 1)
- (10) (j<, A, D, 12)
- (11) (j, _, _, 1)
- (12) (+, A, 2, T_2)
- (13) (:=, T_2 , _, A)
- (14) (j, _, _, 10)

(15) (j, _, _, 1)

【评分标准】

本小题 15 分。

每个四元式 1 分。

如果未能按四元式描述，但描述正确，可酌情给分。

(Optional 2) (15 points)

The following grammar generates binary strings and their complements.

$F \rightarrow B$
 $\quad \mid \neg B$
 $B \rightarrow B0$
 $\quad \mid B1$
 $\quad \mid 0$
 $\quad \mid 1$

The value of a (non-negated) string is just the decimal value of the binary number the string represents; the value of a negated string is the decimal value of the string with 1's replaced by 0's and 0's replaced by 1's. For example, the value of 010 is 2 and $\neg 010$ is 5. Design a syntax-directed definition (SDD) for the above grammar such that the non-terminal F has an attribute $F.val$ which keeps the value of an input string generated by F . Please do NOT modify the grammar.

【参考答案】

SDD 如表 9-1(2)所示

表 9-1(2)

产生式	语义规则
$F \rightarrow B$	$B.c = \text{false}, F.val = B.val$
$F \rightarrow \neg B$	$B.c = \text{true}, F.val = B.val$
$B \rightarrow B_1 0$	$B_1.c = B.c, B.val = B_1.val * 2 + (B.c ? 1 : 0)$
$B \rightarrow B_1 1$	$B_1.c = B.c, B.val = B_1.val * 2 + (B.c ? 0 : 1)$
$B \rightarrow 0$	$B_1.c = B.c, B.val = (B.c ? 1 : 0)$
$B \rightarrow 1$	$B_1.c = B.c, B.val = (B.c ? 0 : 1)$

【评分标准】

本小题 15 分。

10. (10 points) Consider the following basic block:

1)	$S_0 := 2$
2)	$S_1 := 3/S_0$
3)	$S_2 := T - C$
4)	$S_3 := T + C$
5)	$R := S_0/S_3$
6)	$H := R$
7)	$S_4 := 3/S_1$

8)	$S_5 := T + C$
9)	$S_6 := S_4 / S_5$
10)	$H := S_6 * S_2$

- (1) Construct the DAG of the above basic block;
 (2) Assume that only R and H will be used after the basic block. Give the optimized three-address statement sequence.

(1)

【参考答案】

DAG 图:

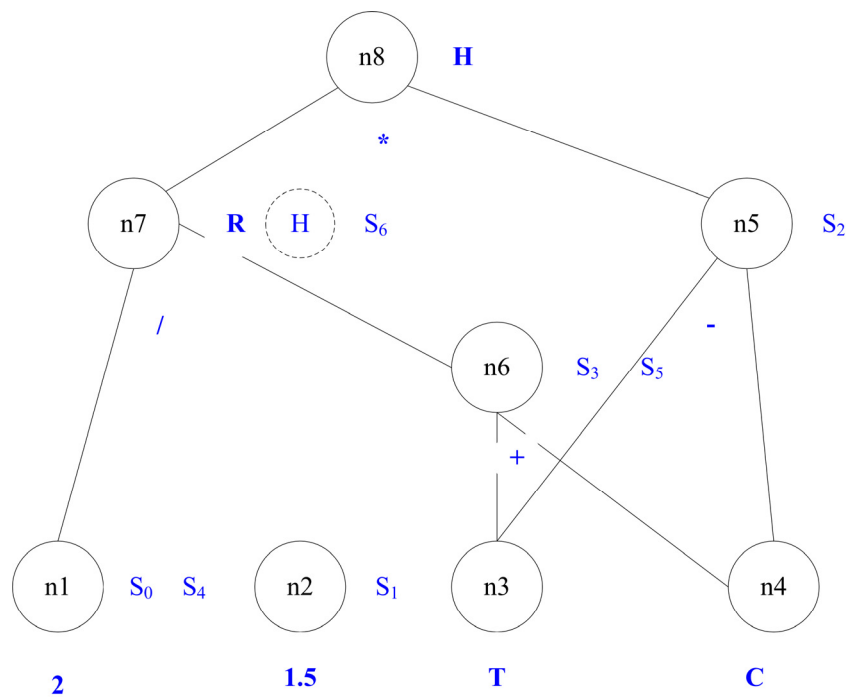


图 10-1

【评分标准】

本小题 5 分。

(2)

【参考答案】

若只有 R,H 在基本块出口活跃, 优化后的四元式序列

$S_2 := T - C$

$S_3 := T + C$

$R := 2 / S_3$

$H := R * S_2$

【评分标准】

本小题 5 分。