

《SE-303 编译原理》期末试题 (B 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 is local optimization?

【参考答案】

局部优化是局限在基本块内的优化。

【评分标准】

本小题 5 分。

2. When shall we do code optimization in a compiler?

【参考答案】

编译的代码优化可在中间代码生成阶段之后，目标代码生成阶段之前进行独立于机器代码优化；还可以在目标代码生成阶段进行依赖于机器的代码优化。

【评分标准】

本小题 5 分。

中间代码生成阶段之后（2 分），目标代码生成阶段之前（1 分）进行独立于机器代码优化（1 分）；还可以在目标代码生成阶段进行依赖于机器的代码优化（1 分）。

3. Give a regular expression for the language over the alphabet $\{a, b\}$: $L = \{a^n b^m \mid (n + m) \text{ is even}\}$

【参考答案】

$(aa)^*(bb)^* \mid a(aa)^*b(bb)^*$

【评分标准】

本小题 5 分。

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

4. (15 points) Given a regular expression as following:

$$(a \mid b)^*(a \mid b)a$$

(1) Based on the *Thompson Algorithm*, construct the NFA from the above regular expression.

(2) Convert the above NFA to a DFA and a minimum-state DFA.

(1)

【参考答案】

根据 *Thompson Algorithm*，对 $(a \mid b)^*(a \mid b)a$ 绘制 NFA 如下图 4-1 所示：

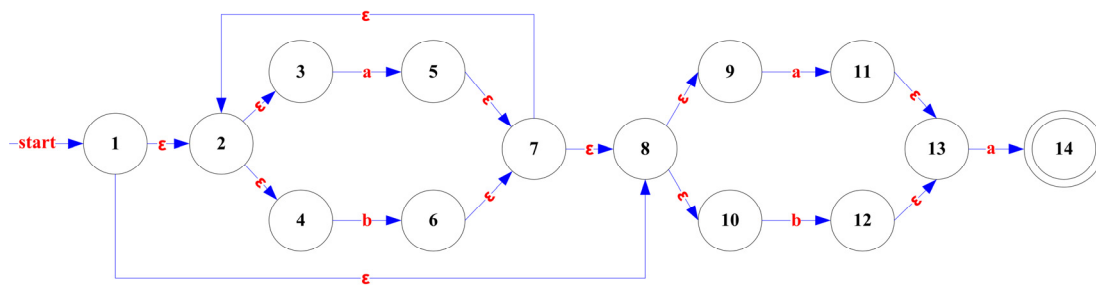


图 4-1

【评分标准】

本小题 6 分。

如图 4-1 所示：(a|b)*对应的 NFA 构造正确得 3 分；(a|b)对应的 NFA 构造正确得 2 分；a 对应的 NFA 构造正确得 1 分。

如果未按 *Thompson Algorithm* 算法绘制，可酌情给分。

(2)

【参考答案】

对图 4-1 的 NFA 先转换为 DFA 如下图 4-2 所示：

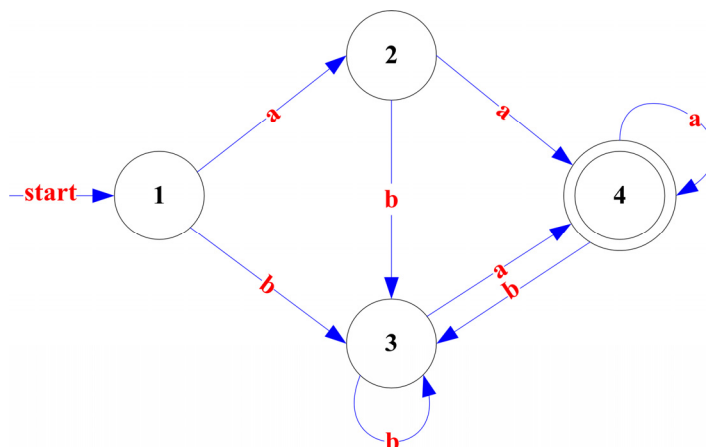


图 4-2

将图 4-2 的 DFA 进行最小化优化，优化后的 DFA 如图 4-3 所示

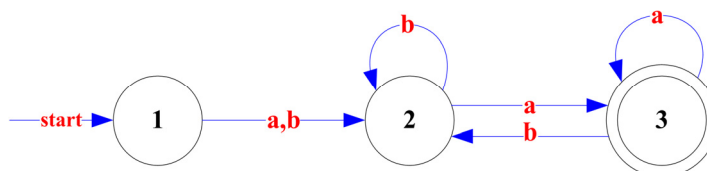


图 4-3

【评分标准】

本小题 9 分。

如图 4-3 所示：优化后的 DFA 正确，但状态数多于 3 或转换关系不完全正确，得 5 分；DFA 正确，状态数等于 3 且状态转换关系正确，得 9 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) $L_1 = \{w \mid w \in \{a,b\}^* \text{ and } w \text{ contain substring } ab\}$;

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

(3) $L_3 = \{w \mid w \in \{a,b\}^* \text{ and } w = manb, m,n \in \{a,b\}^* \mid m| < |n|\}$.

(1)

【参考答案】是正则语言. 正则表达式为: $(a|b)^*ab(a|b)^*$

【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:

$S \rightarrow ACb$

$A \rightarrow BAB \mid a$

$B \rightarrow a \mid b$

$C \rightarrow Ca \mid Cb \mid a \mid b$

【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:

$S \rightarrow SaS \mid SbS \mid ScS \mid d$

(1) Prove this grammar is ambiguous.

(2) Construct an equivalent non-ambiguous grammar for the above grammar.

(1)

【参考答案】

对于此文法的一个句子 $dbdad$, 存在如下图 6-1 所示的两棵分析树, 因此可证明此文法是二义文法。

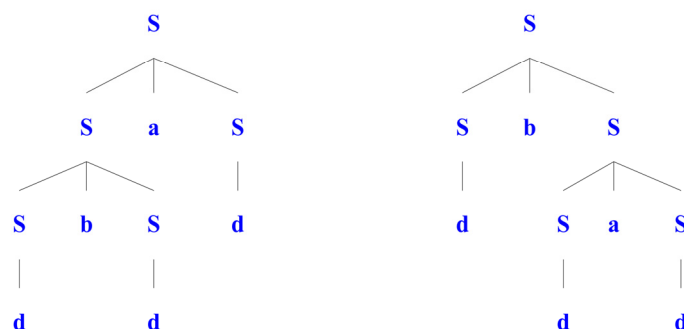


图 6-1

【评分标准】本小题 4 分，正确给出证明过程即可得满分；如果未能给出证明过程，仅给 3 分。

(2)

【参考答案】

$S \rightarrow SaA \mid A$

$A \rightarrow AbC \mid C$

$C \rightarrow CcF \mid F$

$F \rightarrow d$

【评分标准】本小题 6 分，正确给出无二义性文法即可得满分。给出文法与以上文法不同，酌情给分。

7. (10 points) Consider the following grammar:

$S \rightarrow AaAb \mid Bb$

$A \rightarrow \epsilon$

$B \rightarrow \epsilon$

(1) Calculate FIRST and FOLLOW sets for the non-terminals S, A and B.

(2) Construct an LL(1) parsing table for this grammar.

(3) Is this grammar LL(1)? Why?

(1)

【参考答案】

由以上文法消除左递归可得文法：

$FIRST(S) = \{a, b\}$

$FIRST(A) = \{\epsilon\}$

$FIRST(B) = \{\epsilon\}$

$FOLLOW(S) = \{\$ \}$

$FOLLOW(A) = \{a, b\}$

$FOLLOW(B) = \{b\}$

【评分标准】本小题 3 分。

(2)

【参考答案】

预测分析表如表 7-2 所示：

表 7-2

	a	b	\$
S	$S \rightarrow AaAb$	$S \rightarrow Bb$	
A	$A \rightarrow \epsilon$	$A \rightarrow \epsilon$	
B		$B \rightarrow \epsilon$	

【评分标准】本小题 5 分。

(3)

【参考答案】

这一文法是 LL(1)文法，因为 LL(1)分析表中不存在冲突。

【评分标准】本小题 2 分。没有分析原因得 1 分。

8. (10 points) Construct SLR parsing table of the following grammar:

G: $S \rightarrow SS+ \mid SS* \mid a$

【参考答案】

对于文法: $S \rightarrow SS+ \mid SS* \mid a$, 其拓广文法为 G' , 增加产生式 $S' \rightarrow S$, 设产生式排序为:

(0) $S' \rightarrow S$

(1) $S \rightarrow SS+$

(2) $S \rightarrow SS*$

(3) $S \rightarrow a$

【评分标准】

答对此部分可得 2 分。

由产生式知:

$FIRST(S') = FIRST(S) = \{a\}$

$FOLLOW(S') = \{\$ \}$

$FOLLOW(S) = \{*, +, a, \$ \}$

【评分标准】

答对此部分可得 1 分。

G' 的 LR(0) 项目集族及识别活前缀的 DFA 如下图 8-1 所示:

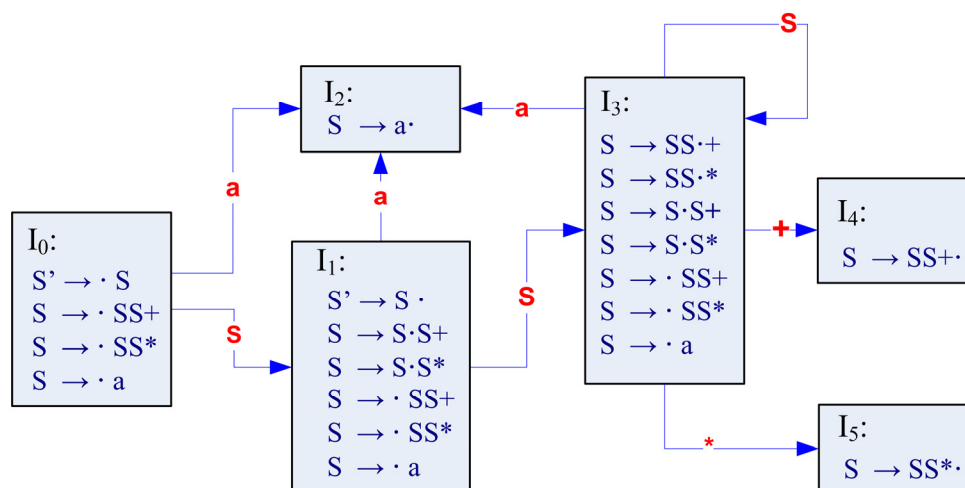


图 8-1

【评分标准】

答对此部分可得 4 分。

SLR 分析表如下表 8-1 所示:

表 8-1

	Action				Goto
	+	*	a	\$	
0			s2		1
1			s2	acc	3
2	r3	r3	r3	r3	
3	s4	s5	s2		3
4	r1	r1	r1	r1	

【评分标准】

答对此部分可得 3 分。

9. (Optional 1) (15 points)

Based on the syntax-directed definitions in the textbook, translate the following program into quadruples, in which $A[i,j]$ is a 20×30 two-dimensional array, the first element is $A[1,1]$.

WHILE (C<D) and (D<E) DO

IF D=10

THEN $A[i,j] := 5 + C$

ELSE $D := E * 2$

【参考答案】

100: if C<D goto 102

101: goto 116

102: if D<E goto 104

103: goto 116

104: if D=10 goto 106

105: goto 113

106: $t1 := i * 30$

107: $t1 := t1 + j$

108: $t2 := A - 124$

109: $t3 := 4 * t1$

110: $t4 := t2[t3]$

111: $t4 := 5 + C$

112: goto 100

113: $t5 := E * 2$

114: $D := t5$

115: goto 100

116:

【评分标准】

本小题 15 分。

每个四元式 1 分。

(Optional 2) (15 points)

The following grammar describe the string that include only '{' and '}' :

$$S \rightarrow T$$

$$T \rightarrow T\{|T\}|\{\}$$

Let S has a synthesized attribute S.valid: when the input string of parentheses is matched, S.valid=true, otherwise, S.valid=false. Design a syntax-directed definition (SDD) for the above grammar to calculate the value of S.valid. If needed, you can create new attributes, but the attributes you create should be a synthesized attribute. Please do NOT modify the grammar. If you want to use the logical operators, please use the ||, && or ! in C++.

(Hint: To match the parentheses, not only the number of parentheses should be the same, but the left or right parentheses should matched form the left side to the right side.eg.} {、} { } { } both are not matched.)

【参考答案】

产生式	语义规则
$S \rightarrow T$	$S.valid = T.valid \ \&\& \ T.count == 0$
$T \rightarrow T_1\{$	$T.valid = T_1.valid$ $T.count = T_1.count + 1$
$T \rightarrow T_1\}$	$T.count = T_1.count - 1$ $T.valid = T_1.valid \ \&\& \ T.count > 0$
$T \rightarrow \{$	$T.valid = true$ $T.count = 1$
$T \rightarrow \}$	$T.valid = false$ $T.count = -1$

【评分标准】

本小题 15 分。

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

1)	$B := 3$
2)	$D := A * C$
3)	$E := A + C$
4)	$F := D - E$
5)	$G := B * F$
6)	$H := A + C$
7)	$I := A * C$
8)	$J := I - H$
9)	$K := B * 5$
10)	$L := K + J$
11)	$M := L$

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

(1)

【参考答案】

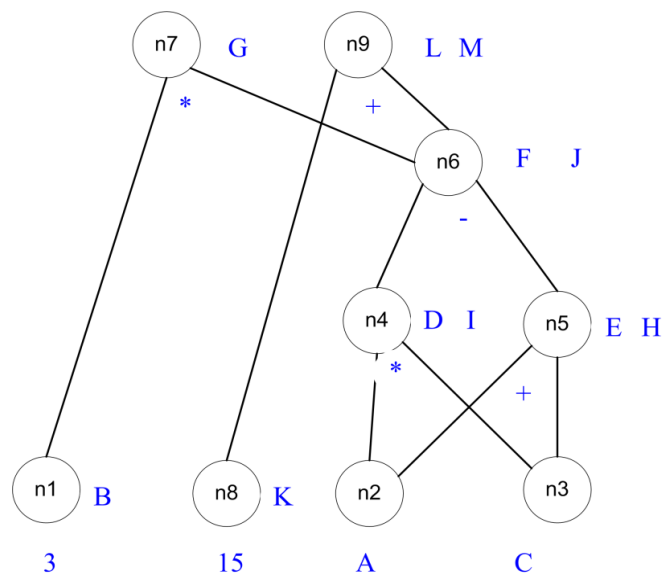


图 10-1

【评分标准】

本小题 5 分。

(2)

【参考答案】

优化后的四元式序列

$T_1 := A * C$

$T_2 := A + C$

$T_3 := T_1 - T_2$

$G := 3 * T_3$

$L := 15 + T_3$

$M := L$

其中: T_1 T_2 T_3 是临时变量。

【评分标准】

本小题 5 分。

注意本题中的临时变量 $T_1 \sim T_3$ 可以替换为其它名称。