

《SE-303 编译原理》期末试题 (A 卷 答案卷)

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

1. Describe some of the tasks that a compiler needs to perform.

【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 Top-Down parsing and Bottom-up parsing?

【参考答案】

自顶向下的语法分析与自底向上的语法分析都可看作为输入串构造语法分析树的问题 (1 分)。

自顶向下的语法分析是从语法分析树的根结点开始, 从顶点开始构造语法树 (2 分)。

自底向上的语法分析是从构成语法分析树叶子结点的终结符号串开始, 从底部开始构造语法分析树 (1 分)。

【评分标准】

本小题 5 分。

其中第一点在后二点中表述也可给分。其它给分细项如上所示。

3. What are S-attributed definition and L-attributed definition?

【参考答案】

S 属性: 对于给定 SDD, 每个属性都是综合属性 (1 分);

L 属性: 对于给定 SDD, 每个属性要么是综合属性 (1 分); 要么是继承属性, 但对于其规则限制如下: 与产生式头部关联的继承属性 (1 分), 位于产生式体部且仅与其左边符号相关的继承属性或综合属性 (1 分), 与这个属性的实例本身相关的继承属性或综合属性 (1 分)。

【评分标准】

本小题 5 分。

S 属性占 1 分;

L 属性占 4 分, 评分细项如上。

Part two: Compute and answer the following questions (5 points for each item.)

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

$$b((ab)^*|bb)^*ab$$

- (1) Based on the Thompson Algorithm, construct the NFA from the above regular expression.
(2) Convert the above NFA to a DFA with minimum number of states.

(1)

【参考答案】

根据 Thompson Algorithm，对 $b((ab)^*|bb)^*ab$ 绘制 NFA 如下图 4-1 所示：

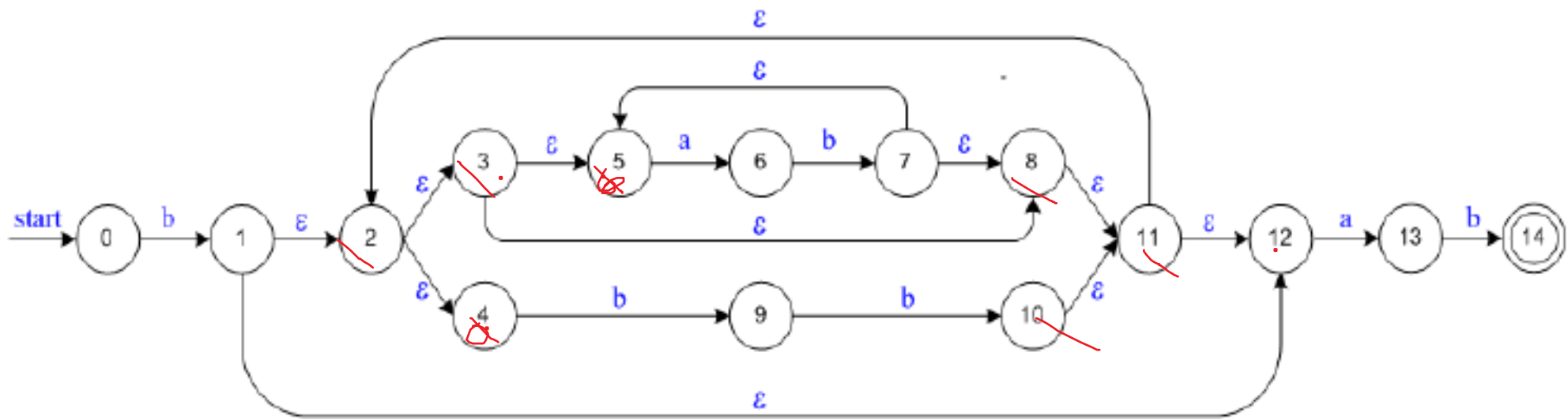


图 4-1

【评分标准】

本小题 6 分。

如图 4-1 所示：(ab)*对应的 NFA 构造正确得 3 分；bb 对应的 NFA 构造正确得 2 分；ab 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，转换过程如表 4-1 所示。

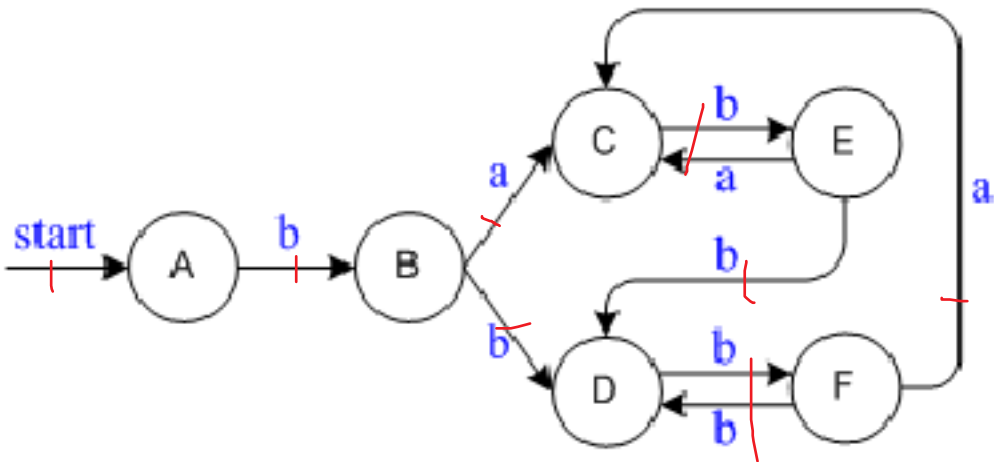


图 4-2

表 4-1

NFA	DFA	A	b
{0}	A	Φ	B
{1,2,3,4,5,8,11,12}	B	C	D
{6,13}	C	Φ	E
{9}	D	Φ	F
{2,3,4,5,7,8,11,12,14}	E	C	D
{2,3,4,5,8,10,11,12}	F	C	D

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。

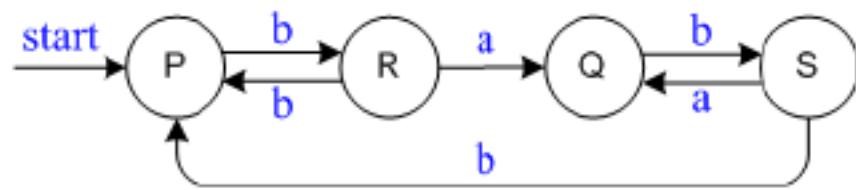


图 4-3

表 4-2

DFA	优化后 DFA 状态
A, D	P
C	Q
B, F	R
E	S

【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 4 或转换关系不完全正确，得 5 分；DFA 正确，状态数等于 4 且状态转换关系正确，得 9 分。如果按以上转换过程进行分析（如图 4-2 所示，表 4-1 及 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) In the following languages, which is a regular language, a context-free language, or not above all? For regular language, write its regular expression, for context-free language (which is not a regular language), write its context-free grammar.

- (1) $L1=\{ a^n b^n | n \geq 1 \}$.
- (2) $L2=\{ a | a \in \{0,1\}^* \text{ and } a \text{ does not contain substring } 10 \}$.
- (3) $L3=\{ a^n b^m c^n d^m | n \geq 1 \wedge m \geq 1 \}$.

(1)

【参考答案】是上下文无关语言，上下文无关方法为： $S \rightarrow aSb \mid ab$ 。

【评分标准】本小题 2 分。

(2)

【参考答案】是正则语言。正则表达式为: 0^*1^* 。

【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。

(3)

【参考答案】是上下文有关语言。

【评分标准】本小题 4 分。答对上下文有关语言得 4 分。（希望同学能说明此题是上下文有关语言的原因）

6. (10 points) Given the following grammar:

- $R \rightarrow R+R$
- $R \rightarrow R \bullet R$
- $R \rightarrow R^*$
- $R \rightarrow (R)$
- $R \rightarrow a \mid b$

(1) Prove this grammar is ambiguous.

(2) Construct an equivalent non-ambiguous grammar for the above grammar, assuming that operator $*$ is higher precedence than \bullet and \bullet is higher precedence than $+$, and all of these operators are left associativity.

(1)

【参考答案】以表达式 $a+b \bullet c$ 进行推导，分别得到推导树如下图 6-1 所示：

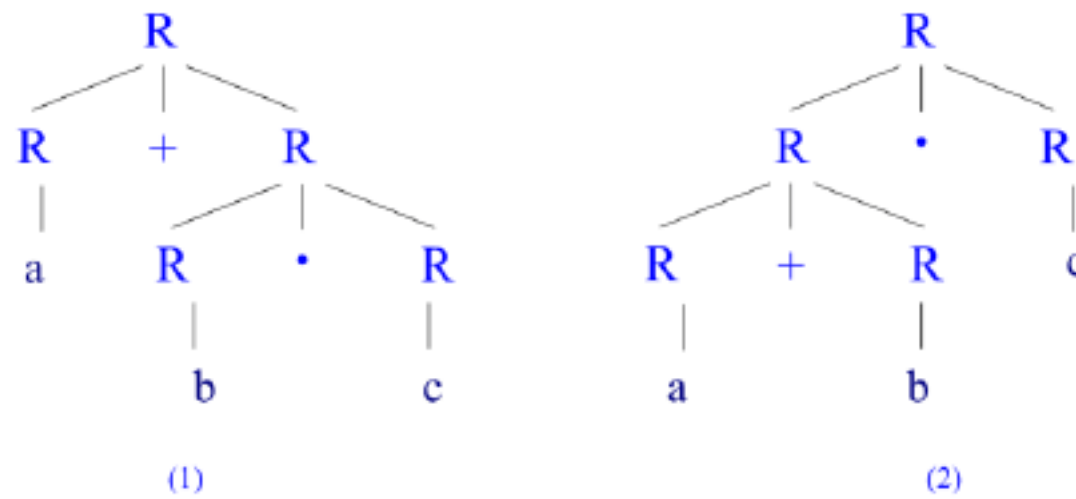


图 6-1

由此可知以上文法具有二义性。

【评分标准】本小题 4 分，正确给出证明过程即可得全分；如果未能给出证明过程，仅给 2 分。

(2)

【参考答案】

构造无二义文法如下：

$R \rightarrow R+T \mid T$
 $T \rightarrow T \bullet F \mid F$
 $F \rightarrow F* \mid P$
 $P \rightarrow (R) \mid a \mid b$

【评分标准】本小题 6 分，正确给出无二义文法即可得全分；给出文法与以上文法不同，酌情给分。

7. (15 points) Consider the following grammar on alphabet $\{ a, o, d, e, f, b \}$:

难

$S \rightarrow MH \mid a$
 $H \rightarrow LSo \mid \epsilon$
 $K \rightarrow dML \mid \epsilon$
 $L \rightarrow eHf$
 $M \rightarrow K \mid bLM$

(1) Write the FIRST and FOLLOW sets for each nonterminals.

(2) Fill in the line of the LL(1) table corresponding to nonterminal S for the grammar.

	a	o	d	e	f	b	\$
S	(1)	(2)	(3)	(4)	(5)	(6)	(7)

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

(1)

【参考答案】

非终结符的 FIRST 集与 FOLLOW 集如下表 8-1 所示:

表 7-1

符号	FIRST 集	FOLLOW 集
S	{a,d,b, ϵ ,e}	{#,o}
M	{d, ϵ ,b}	{e,#,o}
H	{ ϵ ,e}	{#,f,o}
L	{e}	{a,d,b,e,o,#}
K	{d, ϵ }	{e,#,o}

【评分标准】

本小题 6 分，答对 6 个集合即可得满分。

(2)

【参考答案】

- (1) $S \rightarrow a$ ✓
- (2) $S \rightarrow MH$ ✓
- (3) $S \rightarrow MH$ ✓
- (4) $S \rightarrow MH$ ✓
- (5) 无
- (6) $S \rightarrow MH$ ✓
- (7) $S \rightarrow MH$ ✓

【评分标准】

本小题 6 分，答对 6 个产生式即可得满分。

(3)

【参考答案】

此文法是 LL(1)。具体分析如下：

- 对于 $S \rightarrow MH \mid a$, $(FIRST(MH) = \{b, d, e, \epsilon\}) \cap FIRST(a) = \varnothing$,
 $\epsilon \in FIRST(MH)$ 且 $(FOLLOW(S) = \{\$ \}) \cap FIRST(a) = \varnothing$;
 $H \rightarrow LSo \mid \epsilon$, $(FIRST(LSo) = \{e\}) \cap FIRST(\epsilon) = \varnothing$,
 $\epsilon \in FIRST(\epsilon)$ 且 $(FOLLOW(H) = \{f, \$ \}) \cap FIRST(LSo) = \varnothing$;
 $K \rightarrow dML \mid \epsilon$, $(FIRST(dML) = \{d\}) \cap FIRST(\epsilon) = \varnothing$,
 $\epsilon \in FIRST(\epsilon)$ 且 $(FOLLOW(K) = \{e, \$ \}) \cap FIRST(dML) = \varnothing$;
 $M \rightarrow K \mid bLM$, $(FIRST(K) = \{d, \epsilon\}) \cap FIRST(bLM) = \varnothing$,
 $\epsilon \in FIRST(K)$ 且 $(FOLLOW(M) = \{e, \$ \}) \cap FIRST(bLM) = \varnothing$.

综上所述，原文法是 LL(1)文法。

【评分标准】

本小题 3 分。

判断正确得 1 分，解释正确得 2 分。如果给出解释较为简单，酌情给分。

8. (10 points) Answer whether the following grammar is a SLR grammar or not.

- $S \rightarrow L \cdot L \mid L$
 $L \rightarrow LB \mid B$
 $B \rightarrow 0 \mid 1$

【参考答案】

本文法是 SLR 文法，具体分析如下。

(1) 原文法的增广文法如下 G' ：

$S' \rightarrow S$
 $S \rightarrow L \cdot L$
 $S \rightarrow L$
 $L \rightarrow LB$
 $L \rightarrow B$
 $B \rightarrow 0$
 $B \rightarrow 1$

【评分标准】

此部分 2 分。

(2) 非终结符的 FIRST 集与 FOLLO 集如下表 8-1 所示：

表 8-1

	FIRST 集	FOLLO 集
S'	$\{0, 1\}$	$\{\$ \}$
S	$\{0, 1\}$	$\{\$ \}$
L	$\{0, 1\}$	$\{\cdot, 0, 1, \$ \}$
B	$\{0, 1\}$	$\{\cdot, 0, 1, \$ \}$

【评分标准】

此部分 3 分。

(3) 构造增广文法 G' 的 LR(0)项集及可识别活前缀的 DFA 如下图 8-1 所示：

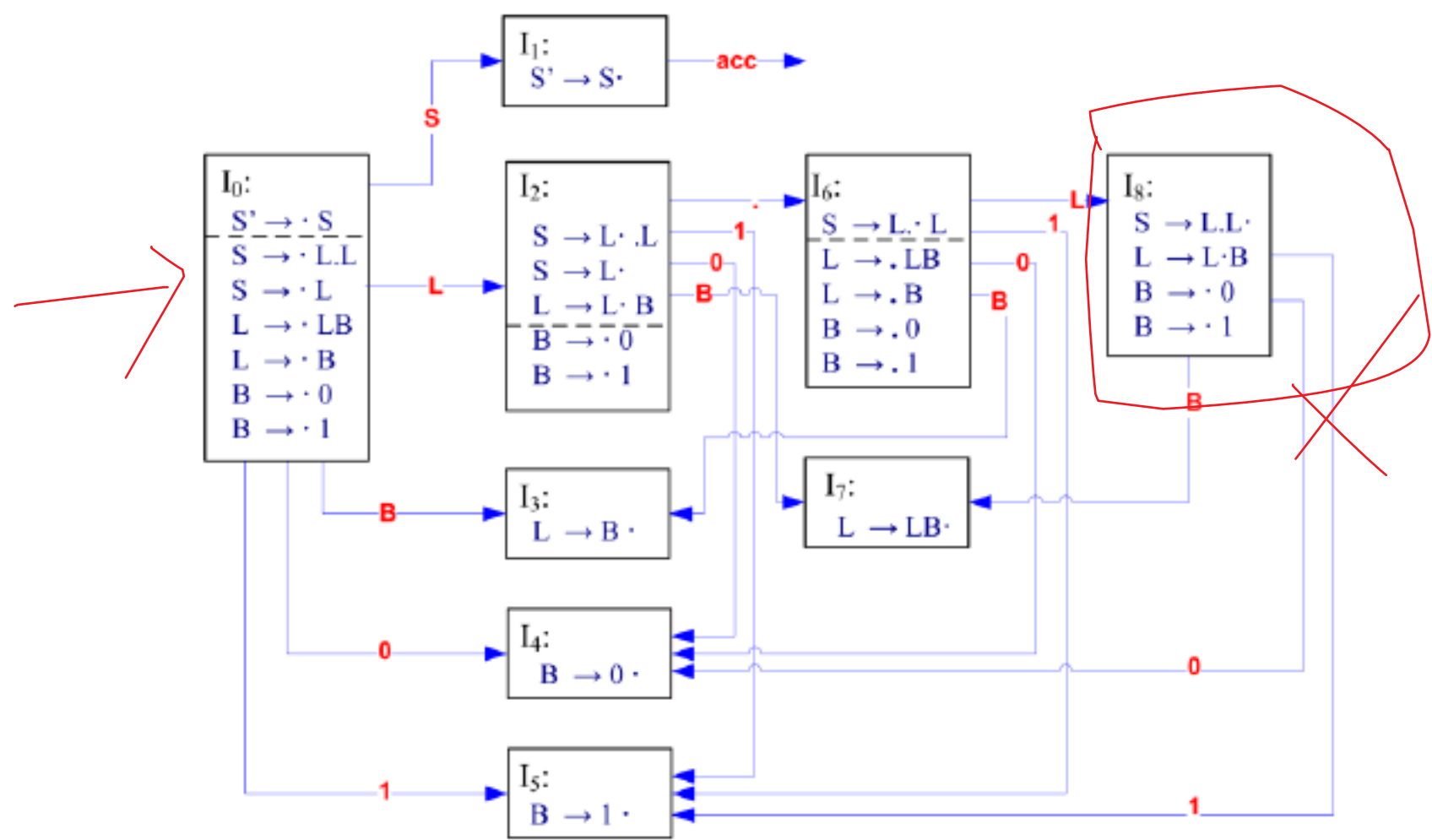


图 8-1

如图 8-1，对于 I_2 ： $B \rightarrow \cdot 0$ 和 $B \rightarrow \cdot 1$ 是移入项目， $S \rightarrow L \cdot$ 为规约项目，存在移入-规约冲突；另外对于 I_8 ， $S \rightarrow L L \cdot$ 也为规约项目，也存在移入-规约冲突，分析过程类似。因此文法不是 LR(0)文法。

【评分标准】

此部分 3 分。如果判断错误，但有完整的分析过程，可酌情给分。

填表实现可能更好

另一方面， I_2 与 I_8 中： $FOLLOW(s) \cap \{0, 1\} = \{\$ \} \cap \{0, 1\} = \varphi$ ，所以 I_2 与 I_8

存在的移入-规约冲突可以由 FOLLOW 集来解决，所以文法 G 是 SLR 文法。

【评分标准】

此部分 2 分。如果判断错误，但有完整的分析过程，可酌情给分。

未

9. (15 points) Based on the syntax-directed definitions in the textbook, translate the following program into quadruples.

```
WHILE (C < D) ∧ (E < D) DO
  IF D = 10 THEN
    WHILE C < E DO C := C+1
  ELSE E := E*2
```

【参考答案】

- (1) (j<, C, D, 3)
- (2) (j, -, -, 0)
- (3) (j<, E, D, 5)
- (4) (j, -, -, 2) ← S.CHAIN
- (5) (j=, D, 10, 7)
- (6) (j, -, -, 13)
- (7) (j<, C, E, 9)
- (8) (j, -, -, 1)
- (9) (+, C, 1, T₁)
- (10) (:=, T₁, -, C)
- (11) (j, -, -, 7)
- (12) (j, -, -, 1)
- (13) (*, E, 2, T₂)
- (14) (:=, T₂, -, E)
- (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。

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

```
A := B + C
B := A - D
C := B + C
D := A - D
```

- (1) Construct the DAG of the above basic block;
- (2) Assume that only variables B are live on exit from the block, produce the optimized three-address codes.
- (3) Assume that only variables B are not live on exit from the block, produce the optimized three-address codes.

(1)

【参考答案】

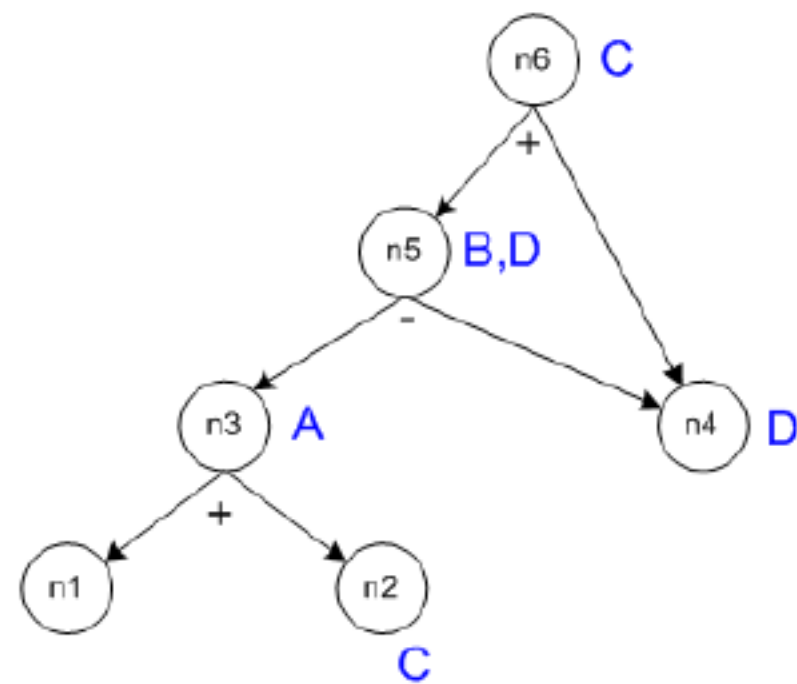


图 10-1

【评分标准】

本小题 4 分。

(2)

【参考答案】

当 B 在出口活跃时，则 B 在出口处要引用，B 的赋值语句就不可删去了，然而 D 与 B 完全一样，故 D 的赋值语句可简化；且考虑到 B 是出口处的唯一活跃变量，可得：

$A := B + C$

$B := A - D$

【评分标准】

本小题 3 分。

(3)

【参考答案】

还是得按照DAG来写

当 B 在出口不活跃时，则 B 在外面就无用，因此 $B := A - D$ 这条赋值语句可删去；另外由于代码生成方面的关系，可把 D 的赋值语句提前到 C 的赋值语句之前。故得到：

$A := B + C$

$D := A - D$

$C := D + C$

【评分标准】

本小题 3 分。