

## “编译原理”考试试卷（A 卷）

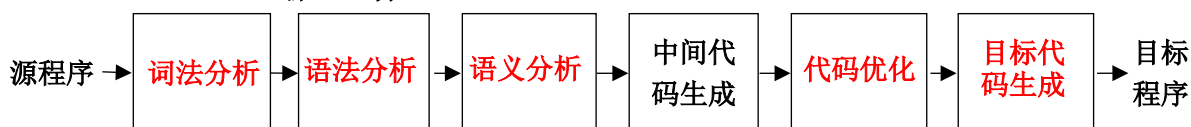
考试方式 闭卷 考试日期 单击或点击此处输入日期。 考试时间 150 分钟  
专业班级                      学 号                      姓 名                     

题号	一	二	三	四	五	六	七	八	九	总分	核对人
分值	10	15	15	10	5	15	15	5	10	100	
得分											

分 数	
评卷人	

### 一、简答（共 10 分）

1. 如果将编译分为六个阶段，请在方框中填写编译阶段的几个主要步骤。（5 分）



2. 请解释什么是编译器的前端和后端，以及这样设计有什么好处。（3 分）

编译程序的前端主要由与源语言有关而与目标机器无关的部分组成，包括词法分析、语法分析、语义分析、中间代码生成、符号表的建立及与机器无关的代码优化工作，前端也包括相应的错误处理工作和符号表操作；

后端由编译程序中与目标机器有关而与源语言无关的部分组成，包括代码优化、目标代码生成，与机器有关的代码优化以及相应的错误处理和符号表操作；

划分前端和后端的好处是便于编译程序的移植和构造。比如重写前端，把一种新的源语言程序编译成同一种中间语言，不用修改后端，即可构造出新源语言在同一目标机器上的编译程序。或者重写后端，可以将同一源语言的编译程序移植到另一种目标机器上。

3. 浅谈使用面向对象方法编写编译器和面向过程方法编写编译器各有什么优势(2 分)；

开放式答题，只要总结合理即可。

解答内容不得超过装订线

分 数	
评卷人	

## 二、文法与语言（共 15 分）

### 1. 给定文法 $G[S]$

$S \rightarrow AB$

$A \rightarrow aA \mid \varepsilon$

$B \rightarrow bc \mid bBc$

请给出其所描述的语言。（5 分）

$$L(G) = \{ a^m b^n c^n \mid m \geq 0, n \geq 1 \}$$

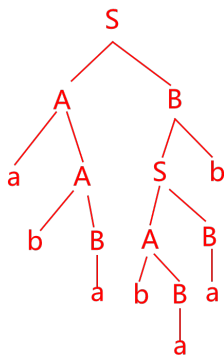
### 2. 设有文法 $G[S]$ :

$S \rightarrow AB$

$A \rightarrow aA \mid bB$

$B \rightarrow a \mid Sb$

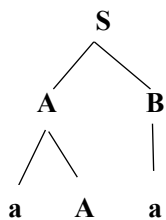
(1) 画出句子  $ababaab$  的语法推导树；（4 分）



(2) 写出句子  $ababaab$  的最左推导；（3 分）

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow abBB \Rightarrow abaB \Rightarrow abaSb \Rightarrow abaABb \Rightarrow ababBBb \Rightarrow ababaBb \Rightarrow ababaab$$

(3) 给定句型  $aAa$  的语法推导树，写出该句型的短语、简单短语和句柄；（3 分）

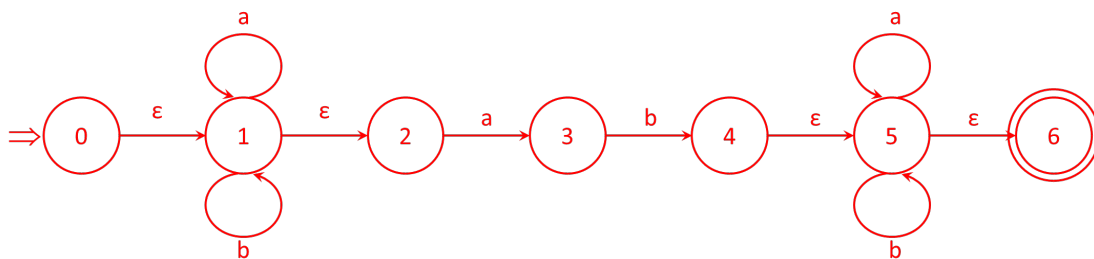


短语:  $aAa, aA, a$ ; 简单短语:  $aA, a$ ; 句柄:  $aA$

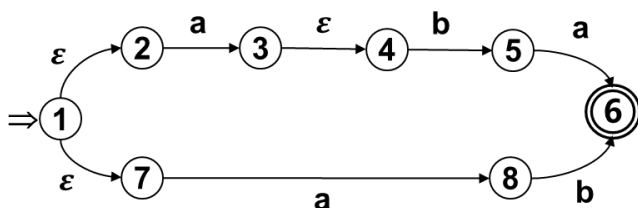
分 数	
评卷人	

### 三、词法分析（共 15 分）

(1) 给定正规式  $(a|b)^*ab(a|b)^*$ ，构造其等价的 NFA；（5 分）



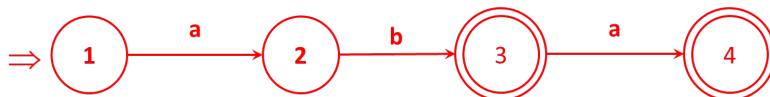
(2) 给定如下 NFA，请将其转换成等价 DFA。（6 分）



子集法：

状态集	a	b
{1, 2, 7}	{3, 4, 8}	-
{3, 4, 8}	-	{5, 6}
{5, 6}	{6}	-
{6}	-	-

重新命名状态集，得 NFA：



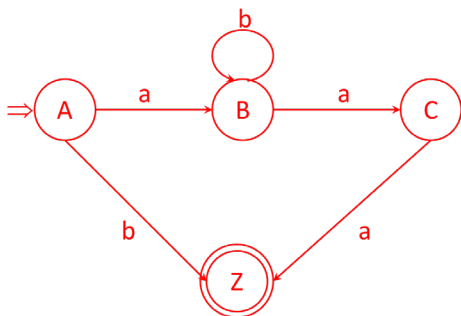
(3) 将下列右线性正规文法  $G[A]$ ：

$A \rightarrow aB | b$

$B \rightarrow aC | bB$

$C \rightarrow a$

转换成等价的 NFA。（4 分）



分 数	
评卷人	

#### 四、自顶向下的语法分析（10 分）

设有文法  $G[S]$ :

$$S \rightarrow TP \quad T \rightarrow +PT \mid \varepsilon \quad P \rightarrow (S) \mid a$$

- (1) 证明  $G[S]$  是 LL(1) 文法（4 分）；
- (2) 构造文法  $G[S]$  的 LL(1) 预测分析表；（3 分）
- (3) 根据 LL(1) 分析表，分析符号串 (a) 是否该文法的句子(注意：符号串(a)中，括号为符号串的一部分)。列表写出分析过程，列表内容包括：步聚、符号栈内容、待分析串内容、每步的动作或结论。（3 分）

(1)  $\text{Select}(S \rightarrow TP) = \{+, (, a\}$ ;

$\text{Select}(T \rightarrow +PT) = \{+\}$ ;  $\text{select}(T \rightarrow \varepsilon) = \{ (, a \}$ 。显有  $\text{Select}(T \rightarrow +PT) \cap \text{select}(T \rightarrow \varepsilon) = \phi$ ;

$\text{Select}(P \rightarrow (S)) = \{ ( \}$ ;  $\text{select}(P \rightarrow a) = \{ a \}$ 。显有  $\text{Select}(P \rightarrow (S)) \cap \text{select}(P \rightarrow a) = \phi$ ;

因为每个非终结符的不同产生式的 SELECT 集无交集，故文法  $G[S]$  为 LL(1) 文法。

(2)  $G[S]$  的预测分析表：

	+	a	(	)
S	$\rightarrow TP$	$\rightarrow TP$	$\rightarrow TP$	
T	$\rightarrow +PT$	$\rightarrow \varepsilon$	$\rightarrow \varepsilon$	
P		$\rightarrow a$	$\rightarrow (S)$	

(2) 对符号 (a) 的分析过程：

步骤	符号栈	输入串	产生式或匹配
1	#S	(a)#	$S \rightarrow TP$
2	#PT	(a)#	$T \rightarrow \varepsilon$
3	#P	(a)#	$P \rightarrow (S)$
4	#)S(	(a)#	‘(’ 匹配
5	#)S	a)#	$S \rightarrow TP$
6	#)PT	a)#	$T \rightarrow \varepsilon$
7	#)P	a)#	$P \rightarrow a$
8	#)a	a)#	‘a’ 匹配
9	#)	)#	‘)’ 匹配
10	#	#	接受

分 数	
评卷人	

五、(共 5 分) 设有文法  $G[S]$ :

$S \rightarrow D(R) \quad R \rightarrow R;P|P \quad P \rightarrow S|i \quad D \rightarrow i$

求  $G$  中每个非终结符的 FIRSTVT 集和 LASTVT 集;

FIRSTVT(S) = { (, i }	LASTVT(S) = { ) }
FIRSTVT(R) = { ;, (, i }	LASTVT(R) = { ;, ), i }
FIRSTVT(P) = { (, i }	LASTVT(P) = { i, ) }
FIRSTVT(D) = { i }	LASTVT(D) = { i }

分 数	
评卷人	

六、自底向上的语法分析 (共 15 分)。

设有文法  $G[S]$ :

$S \rightarrow Aa|bAc|dc \quad A \rightarrow d$

(1) 拓广文法, 构造 LR(1) 项目集规范族及识别活前缀的 DFA; (6 分)

(2) 构造该文法的 LR(1) 分析表 (6 分);

(3) 用表格给出输入串  $bdc$  的 LR(1) 分析过程, 包括步骤、状态栈、符号栈、输入串的变化, 以及 ACTION 和 GOTO 等动作。(3 分)。

答: (1) 先拓展文法:

(0)  $S' \rightarrow S$

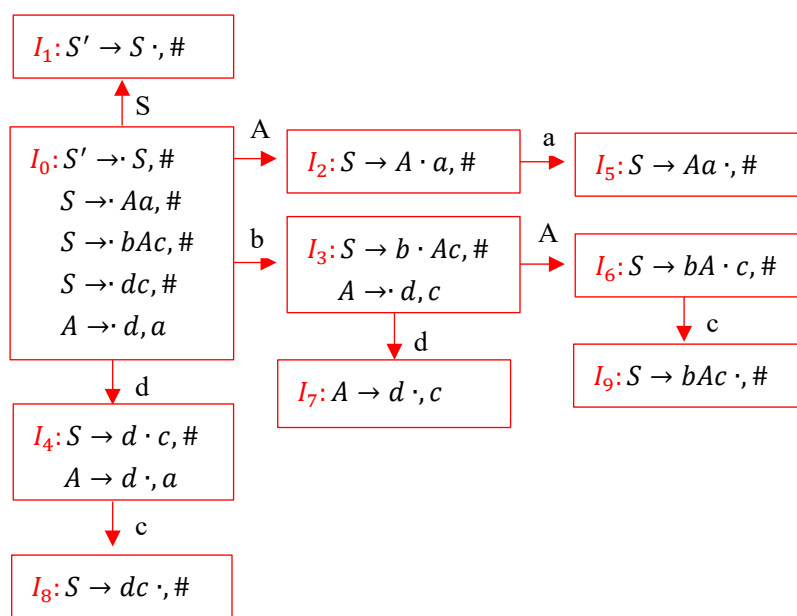
(1)  $S \rightarrow Aa$

(2)  $S \rightarrow bAc$

(3)  $S \rightarrow dc$

(4)  $A \rightarrow d$

再构造 LR(1) 自动机:



(2) 由 DFA 构造 LR(1)分析表

	ACTION					GOTO	
	a	b	c	d	#	S	A
0		S3		S4		1	2
1					ACC		
2	S5						
3				S7			6
4	r4			S8			
5					r1		
6			S9				
7			r4				
8					r3		
9					r2		

(3) 输入串 bdc 的分析过程:

步骤	状态栈	符号栈	输入串	ACTION	GOTO
1	0	#	bdc#	S3	
2	03	#b	dc#	S7	
3	037	#bd	c#	r4:A→d	6
4	036	#bA	c#	S9	
5	0369	#bAc	#	r2: S→bAc	1
6	01	#S	#	ACC	

分 数	
评卷人	

七、语法制导的翻译模式及中间代码生成（共 15 分）。

下面是某语言文法的部分产生式及相应的语义动作集合：

$S \rightarrow id = E \text{ “;” } \{ S.code := E.code \parallel gen(id.place \text{ ‘:=’ } E.place) \}$   
 $S \rightarrow \text{while } ( \{ E.true := newlabel; E.false := S.next \} E ) \text{ “{”}$   
 $\quad \{ S.next := newlabel \} S_1 \text{ “}”$   
 $\quad \{ S.code := gen(S_1.next \text{ ‘:’ } ) \parallel E.code \parallel gen(E.true \text{ ‘:’ } ) \parallel S_1.code$   
 $\quad \parallel gen( \text{ ‘goto’ } S_1.next ) \}$   
 $S \rightarrow \text{if } ( \{ E.true := newlabel; E.false := newlabel \} E )$   
 $\quad \text{“{” } \{ S_1.next := S.next \} S_1 \text{ “}” \text{ else “{” } \{ S_2.next := S.next \} S_2 \text{ “}”}$   
 $\quad \{ S.code := E.code \parallel gen(E.true \text{ ‘:’ } ) \parallel S_1.code \parallel gen( \text{ ‘goto’ } S.next ) \parallel$   
 $\quad gen(E.false \text{ ‘:’ } ) \parallel S_2.code \}$   
 $E \rightarrow \text{num } \{ E.place := newtemp; E.code := gen( E.place \text{ ‘:=’ } num.val ) \}$   
 $E \rightarrow id \{ E.place := id.place; E.code := \text{ “” } \}$   
 $E \rightarrow E_1 + E_2 \{ E.place := newtemp; E.code := E_1.code \parallel E_2.code$   
 $\quad \parallel gen( E.place \text{ ‘:=’ } E_1.place \text{ ‘+’ } E_2.place ) \}$   
 $E \rightarrow id_1 \text{ rop } id_2$   
 $\quad \{ E.code := gen( \text{ ‘if’ } id_1.place \text{ rop. op } id_2.place \text{ ‘goto’ } E.true )$   
 $\quad \parallel gen( \text{ ‘goto’ } E.false ) \}$   
 $E \rightarrow ( \{ E_1.true := E.true; E_1.false := E.false \} E_1 )$   
 $\quad \{ E.code := E_1.code \}$

语义属性说明：

id.place : id 对应的符号表的存储位置；

E.place : 用来存放 E 的值的存储位置；

E.code : 对 E 求值的三地址代码序列；

E.true 和 E.false 分别表示布尔表达式为真和假时，程序转移的目标位置。

S.code : 对应于 S 的三地址代码序列；

S.next: 表示 S 之后要执行的首条 TAC 语句的标号。

语义函数/过程说明：

gen() : 生成一条三地址代码；

newtemp : 在符号表中新建一个从未使用过的名字, 并返回该名字的存储位置；

||: 是三地址代码序列之间的链接运算；

newlabel 返回一个新的语句标号。

(1) 上述语义规则和动作采用的是什么类型的翻译模式；（1 分）

(2) 根据题设给出的文法画出语句 S0:

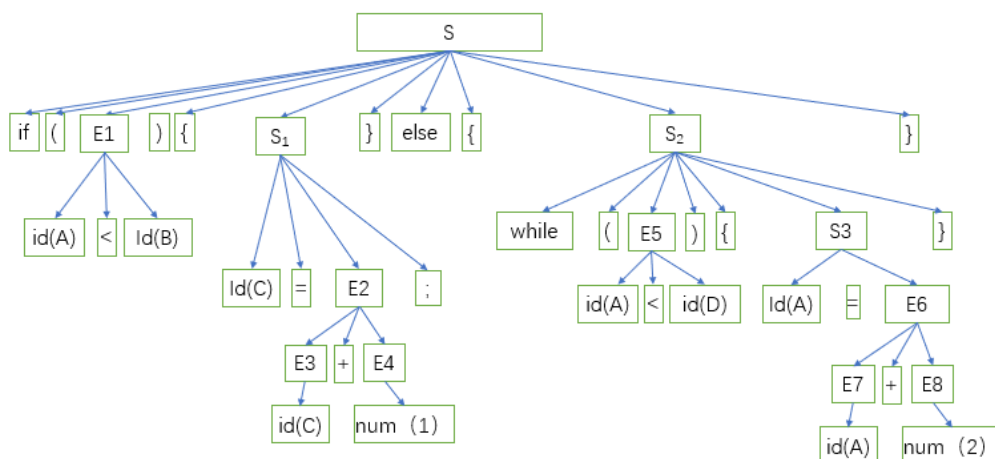
$\text{if } (a < b) \{ c = c + 1; \} \text{ else } \{ \text{while } (a < d) \{ a = a + 2; \} \}$

的语法分析树；（6 分）

(3) 根据题设给出的文法和语义规则, 计算语法树各结点的继承属性值, 并将语句 S0 翻译成三地址代码序列。（8 分）

答：(1) L 翻译模式；（1 分）

(2) 语法树见下图：



(共 6 分) 树第一层 (2 分), while 那一层 (1 分), A<B (1 分), C=C+1(1 分), A=A+2 (1 分)

(3) 继承属性: E.true, E.false, S.next , 假设 S.next = 10

E1.true=11;      E1.false=12

S1.next = 10;      S2.next = 10

E5.true=13;      E5.false=10

S3.next = 14;

if a<b goto 11

goto 12

11: temp1:= 1

temp2 := c+temp1

c:=temp2

goto 10

12:

14: if (a<d) goto 13

goto 10

13: temp3:= 2

temp4 := a+temp3

a:=temp4

goto 14

10:



分 数	
评卷人	


八、运行时存储组织（5 分）。

现有 c 语言的程序片段：

```
void f(int i1, int i2)
{
    int i3;
    int i4[2][2];
    i3 = 7;
    i4[0][0] = 1;
    i4[0][1] = 9;
    i4[1][0] = 8;
    i4[1][1] = 5;
    /*程序运行点1*/
}

int main()
{
    f(3, 4);
    return 0;
}
```

假设某编译器给出的栈帧结构如下：

实际参数		栈生长方向
返回地址		
相关寄存器		
局部变量区		

假设该编译器没有开启任何优化选项，并且对于函数调用时，参数是从右向左依次入栈, 临时变量使用寄存器存储。程序经过该编译器编译成目标代码后在某平台上运行，当运行到“程序运行点 1”时，请填写函数 f 的栈帧内容。

i2(4)
i1(3) (方向 1 分，值 1 分)
返回地址
相关寄存器
i3(7) (1 分)
I4[0][1](5)
I4[1][0](8)
I4[0][1](9)
I4[0][0](1) (方向 1 分，值 1 分)

或

i2(4)
i1(3) (方向 1 分，值 1 分)
返回地址
相关寄存器
I4[0][1](5)
I4[1][0](8)
I4[0][1](9)
I4[0][0](1) (方向 1 分，值 1 分)
i3(7) (1 分)

分 数	
评卷人	

九、代码优化（10 分）。

有如下三地址代码片段，前面 L1~L17 是标号：

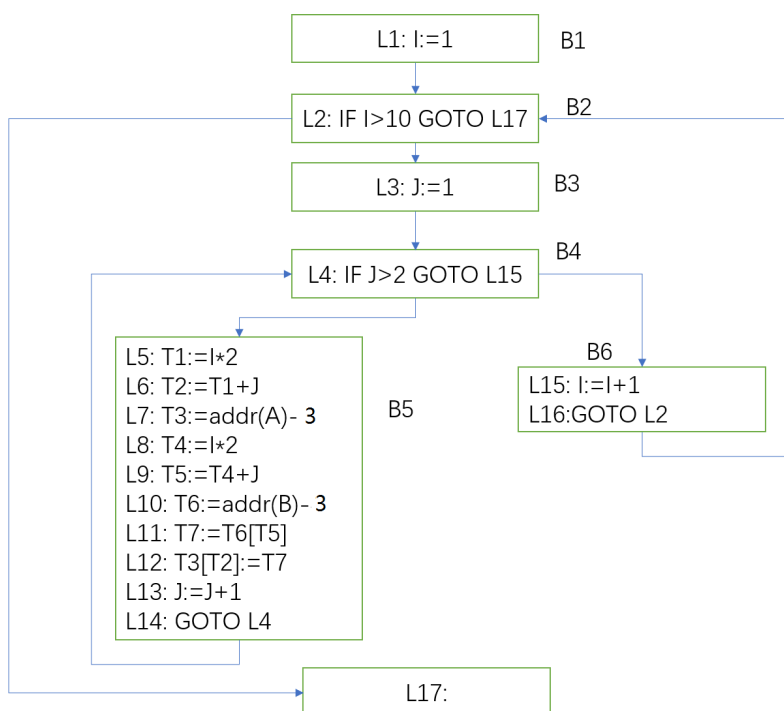
```

L1: I:=1
L2: IF I>10 GOTO L17
L3: J:=1
L4: IF J>2 GOTO L15
L5: T1:=I*2
L6: T2:=T1+J
L7: T3:=addr(A)-3
L8: T4:=I*2
L9: T5:=T4+J
L10: T6:=addr(B)-3
L11: T7:=T6[T5]
L12: T3[T2]:=T7
L13: J:=J+1
L14: GOTO L4
L15: I:=I+1
L16: GOTO L2
L17:

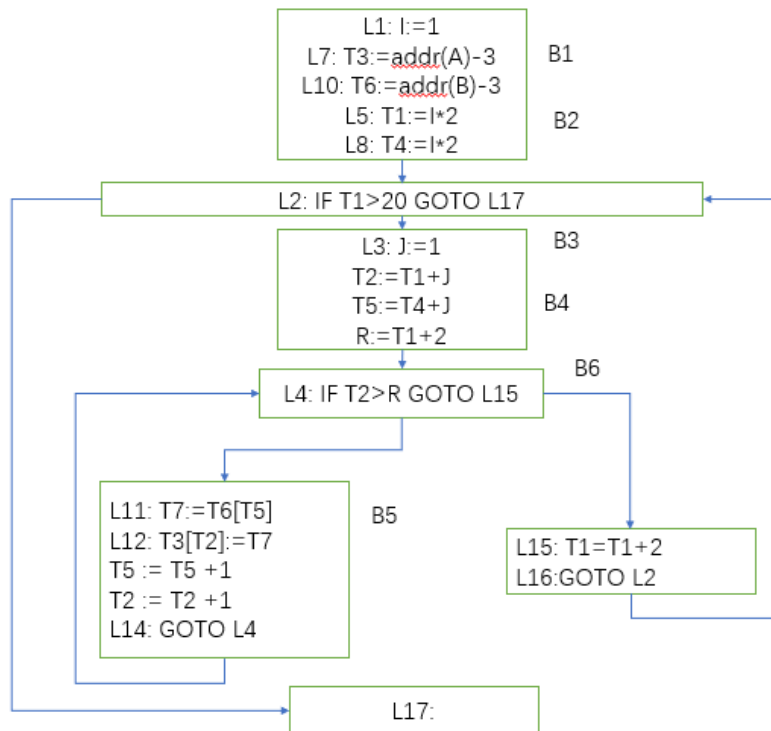
```

- (1) 请将三地址码序列划分为基本块并给出流图(3 分)；
- (2) 找出流图中的循环 (3 分)；
- (3) 对上述代码进行代码外提、删除归纳变量、强度削弱及其它优化;(4 分)；

答：(1)基本块划分及流图：



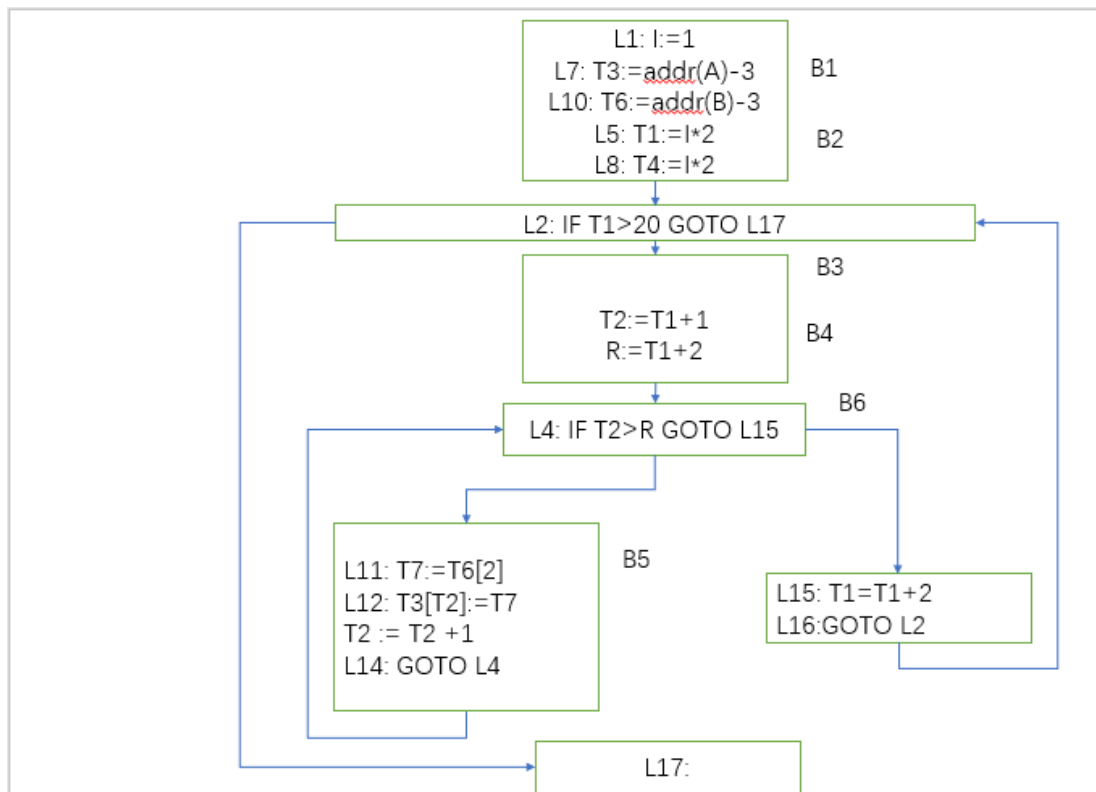
(2) 循环: {B4, B5}; {B2, B3, B4, B5, B6}



归纳变量: I、J 改为 T1 和 T2 控制, 即数组索引 (1 分)

代码外提: T3=addr (A) -3, T6=addr (B) -3 (1 分)

强度削弱: T5:=T5+1, T2:=T2+1 (1 分)



经过删除多余运算、常量合并等优化后的结果 (1 分)