

一、简答题（每题各 5 分，共 10 分）

1、编译程序的前端与后端是如何划分的，各包括哪些部分，这样的划分的好处是什么？

编译前端主要包括词法分析、语法分析、语义分析、中间代码生成这几个部分，后端则包含代码优化和目标代码生成部分。前端的特点是仅与编译的源语言有关，而后端则仅与编译的目标语言及运行环境有关。将编译过程划分成前端和后端，主要目的是在多种源语言和多种目标语言的开发过程中，可以灵活搭配组合，消除重复开发的工作量，提高编译系统的开发效率。

2、已知文法 $G[S]$ 是算符优先文法： $S \rightarrow S \mid F \mid F$ $F \rightarrow F \mid B \mid B$ $B \rightarrow a \mid b$ 构造每个非终结符号的 FIRSTVT 与 LASTVT 集。

$$\text{FirstVT}(\text{S}) = \{a, b, m, n\}; \text{FirstVT}(\text{F}) = \{a, b, n\}; \text{FirstVT}(\text{B}) = \{a, b\}$$

$$\text{LastVT}(\mathbf{B}) = \{a, b\}; \text{LastVT}(\mathbf{F}) = \{a, b, n\}; \text{LastVT}(\mathbf{S}) = \{a, b, m, n\};$$

二. 写出一个文法使其语言为: $L = \{a^n b^m a^{2n} \mid n, m \geq 0\}$ (5分)

$$S \rightarrow aSaa \mid B$$

$$B \rightarrow bB \mid \varepsilon$$

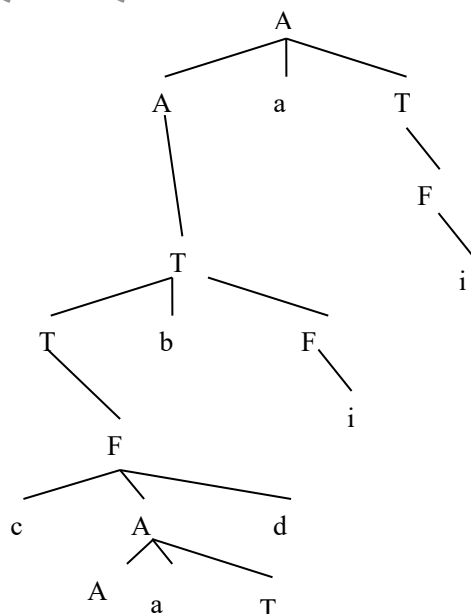
三. 已知文法 $G[A]$: (13 分)

$$A \rightarrow AaT \mid T \quad T \rightarrow TbF \mid F \quad F \rightarrow cAd \mid i$$

(1) 给出句型 cAaTdbiai 最右推导, 构造其语法分析树。

(2) 写出句型的所有短语、直接短语、句柄和素短语。

(1) $A \Rightarrow AaT \Rightarrow AaF \Rightarrow AaI \Rightarrow Tai \Rightarrow i = TbFai \Rightarrow Tbiai \Rightarrow Fbiai \Rightarrow cAdbia i \Rightarrow cAaTdbia i$



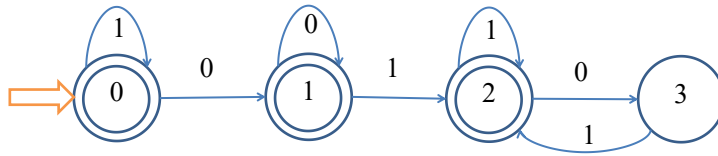
(2) 短语: i, AaT, cAaTd, cAaTdbi, cAaTdbiai

直接短语: AaT, i

句柄: AaT

素短语: AaT,i

四. 已知正规式 $R=1^*(0^*1)(011^*)^*$ 给出与之等价的最小 DFA 。 (15 分)



五. 已知文法 $G(S)$: (15 分)

$S \rightarrow S B$ $S \rightarrow y$ $B \rightarrow B x$ $B \rightarrow A x$
 $A \rightarrow z$ $A \rightarrow z S y$

- (1) 将其消除左递归并提取公共左因子;
- (2) 求改造后文法的每一个非终结符的 First 集合与 Follow 集合;
- (3) 给出改造后文法的预测分析表。

(1) 消除左递归, 提取公共左因子。

$S \rightarrow y S'$ $S' \rightarrow B S'$ $S' \rightarrow \epsilon$
 $B \rightarrow A x B'$ $B' \rightarrow x B'$ $B' \rightarrow \epsilon$
 $A \rightarrow z A'$ $A' \rightarrow S y$ $A' \rightarrow \epsilon$

$First(S) = \{y\}$ $First(S') = \{\epsilon, z\}$ $First(B') = \{x, \epsilon\}$ $First(A) = \{z\}$ $First(A') = \{y, \epsilon\}$ $First(B) = \{z\}$

$Follow(S) = \{ \#, y \}$ $Follow(S') = \{ \#, y \}$ $Follow(B) = \{z, \#\}$ $Follow(A) = \{x\}$
 $Follow(B') = \{ \#, z, y \}$ $Follow(A') = \{x\}$

Follow(S)={ #, y } Follow(S')={ #, y } Follow(B)={z, #} Follow(A)={x}

Follow(B')={ #,z,y } Follow(A')={x}

	y	x	z	#
S	$S \rightarrow yS'$			
S'	$S' \rightarrow \epsilon$		$S' \rightarrow BS'$	$S' \rightarrow \epsilon$
B			$B \rightarrow Ax B'$	
B'	$B' \rightarrow \epsilon$	$B' \rightarrow x B'$	$B' \rightarrow \epsilon$	$B' \rightarrow \epsilon$
A			$A \rightarrow z A'$	
A'	$A' \rightarrow S y$	$A' \rightarrow \epsilon$		

预测分析表不含有多重入口, 是 LL (1) 文法

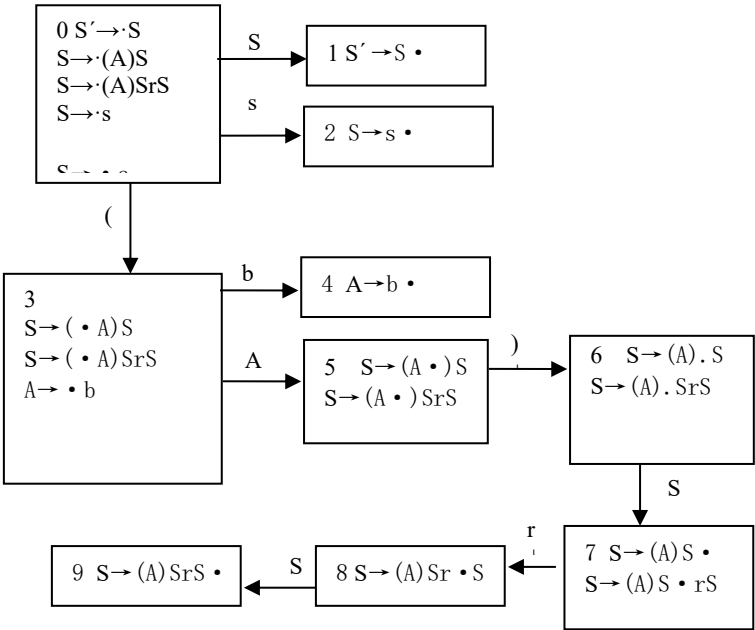
六. 设有文法 $G(S)$: (15 分)

$S \rightarrow (A) S \mid (A) S r S \mid s$

$A \rightarrow b$

- (1) 构造识别该文法所有活前缀的项目集规范簇和 DFA。
- (2) 构造其 SLR 分析表, 判断其是否为 SLR (1) 文法。

(0) $S' \rightarrow S$ (1) $S \rightarrow (A) S$ (2) $S \rightarrow (A) S r S$ (3) $S \rightarrow s$ (4) $A \rightarrow b$

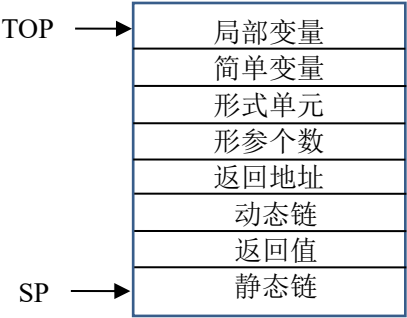


Follow(S) = {r, #}
Follow(A) = {) }

状态	ACTION						GOTO	
	()	r	s	b	#	S	A
0	s3			S2			1	
1						✓		
2			r3			r3		
3					s4		5	
4		r4						
5		s6						
6							7	
7			r1/s8			r1		
8							9	
9			r2			r2		

由于分析表含多重入口，所以该文法不是 SLR (1) 文法。

七. (1)已知某语言过程调用、进入、数据空间分配和过程返回翻译过程如下：(7 分)



```

par  Ti {
    ( i+4 ) [TOP] := Ti      (传递参数值)
    或者 ( i+4 ) [TOP] := addr(Ti) (传递参数地址)
}

```

```

Call P, n {
    (说明: P 为过程名, n 为参数个数)
    ( 3[TOP] ) := SP
    5[TOP]      := n
    ( 1[TOP] ) := 静态链
    JSR P
}

```

```

过程进入后: {
    SP := ( TOP+1 )
    ( 3[SP] ) := 返回地址
    TOP := TOP + L
}

```

```

Return {
    TOP := ( SP-1 )
    SP := ( 3[SP] )
    X := 4[TOP]
    UJ 0[X]
}

```

(2) 某语言描述的是十进制是4的倍数的二进制串。对应的文法为: (8分)

$S \rightarrow L100 \quad L \rightarrow LB \quad L \rightarrow B \quad B \rightarrow 0 \quad B \rightarrow 1$

请给出求解该二进制串对应的十进制的翻译模式, 并概述其基本设计思想。

主要设计思想是用继承属性pos记录当前 2^n 中n的值。Value属性是综合属性, 记录每累加的结果。

$S \rightarrow \{L.pos = 3\} L100 \{S.value = L.value + 4\}$

$L \rightarrow \{L.pos = L.pos + 1\} L1 \{B.pos = L.pos\} B \{L.value = L1.value + B.value;\}$

$L \rightarrow \{B.pos = L.pos\} B \{L.value = B.value\}$

$B \rightarrow 0 \quad \{B.value = 0\}$

$B \rightarrow 1 \quad \{B.value = 2^{b.pos}\}$

八. 有如下的代码块: (12分)

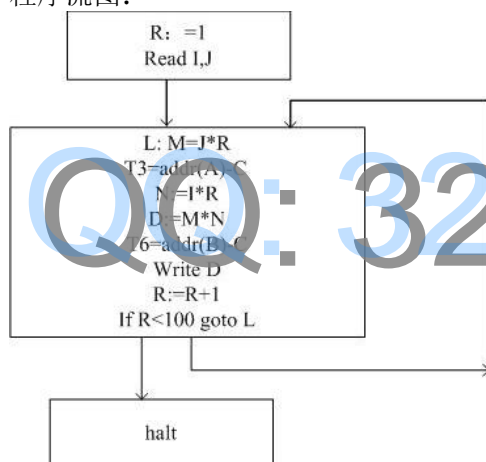
```

R:=1
read I,J
L: M:=J*R
T3=addr(A)-C
N:=I*R
D:=M*N
T6=addr(B)-C
write D
R:=R+1
If R<100 goto L
halt

```

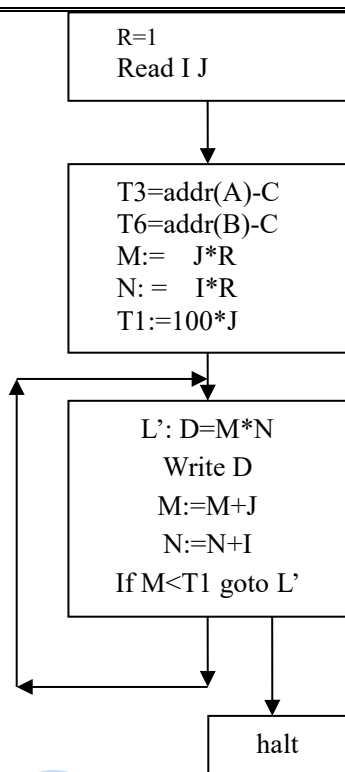
- (1) 划分基本块，并画出程序流图；
- (2) 进行最大可能的循环优化

程序流图：



先进行基本块划分，再画程序流图。先将该程序划分为 3 个基本块，程序流图如图所示。

- (1) 代码外提。T3:=addr(A)-C T6:=addr(B)-C
- (2) 强度削弱。由于循环中 $M := J * R$, $N := I * R$ ，其中 I/J 在循环中值不发生改变， I 每次增加 1。因此对 A, B 的赋值运算可进行强度削弱。即将表达式中的乘法运算改为加法运算。
- (3) 删除基本归纳变量。循环中 R 是基本归纳变量， M, N 是与 R 同族的归纳变量。 $M = J * R$, $N = I * R$ 。因此，条件 $R < 100$ ，可用 $M < 100 * J$ 或 $N < 100 * I$ 替代。这样控制条件可改为： $T1 := 100 * J$ if $M < T1$ goto L' 或者 $T2 := 100 * I$ if $N < T2$, goto L'。控制条件经过以上改变之后，就可以删除基本快中的语句 $R := R + 1$ 。又语句 $T1 := 100 * J$ 是循环中的不变运算。可以代码外提。结果如图所示。



一. 简答题

(1) 给出编译程序的基本结构, 并给出编译器常见的几种中间语言。

编译程序主要包括词法分析、语法分析、语义分析、中间代码生成、代码优化和目标代码生成等部分。

常见的几种中间语言: 后缀式、三地址代码 (三元式、四元式、间接三元式)、DAG 图等

(2) 已知文法 $G(S): S \rightarrow bTc \quad S \rightarrow a \quad T \rightarrow R \quad R \rightarrow R/S \quad R \rightarrow S$, 请给出每一个非终结符的 FIRSTVT 和 LASTVT 集合。

$\text{FIRSTVT}(S) = \{a, b\} \quad \text{FIRSTVT}(T) = \{a, b, / \} \quad \text{FIRSTVT}(R) = \{a, b, / \}$

$\text{LASTVT}(S) = \{a, c\} \quad \text{LASTVT}(T) = \{a, c, / \} \quad \text{LASTVT}(R) = \{a, c, / \}$

(3) 运行时存储空间划分主要包括哪些区域?

目标代码区, 静态数据区, 栈区, 堆区

二. 写一文法使其语言为: $L(G) = \{a^m b^n c^n \mid m \text{ 为奇数, } n \text{ 为偶数且 } n > 0\}$

$S \rightarrow AC$

$A \rightarrow aaA|a$

$B \rightarrow bbBcc \mid bbcc$

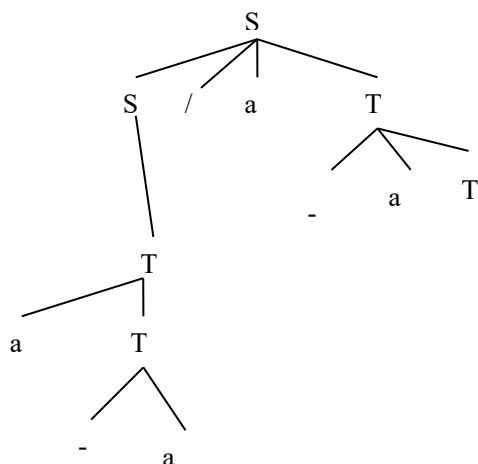
三. 已知文法 $G(S)$:

$S \rightarrow S/aT \mid aT \mid /aT$

$T \rightarrow -aT \mid -a$

(1) 请给出句型 $a-a/a-aT$ 的最左推导, 并画出其语法分析树。

$S \Rightarrow S/aT \Rightarrow aT/aT \Rightarrow a-a/aT \Rightarrow a-a/a-aT$



(2) 写出上述句型的所有短语、直接短语、句柄和素短语。

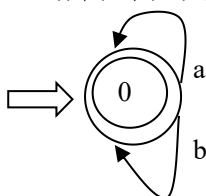
短语: $-a, a-a, -aT, a-a/a-aT$

直接短语: $-a, -aT$

句柄: $-a$

最左素短语: $-a$

四. 已知正规式表达式 $r=((a|b)^*|(bb)^*)^*$ 给出与之等价的最小 DFA。



五. 已知文法 $G(E)$:

$E \rightarrow [F]E \mid [F]$

$F \rightarrow i \mid Fi$

(1) 通过消除左递归和提取左公因子, 将其改造成 LL(1) 文法;

消除左递归:

$$E \rightarrow [F]E \mid [F]$$

$$F \rightarrow iF'$$

$$F' \rightarrow iF' \mid \varepsilon$$

提取左公因子:

$$E \rightarrow [F]E'$$

$$E' \rightarrow E \mid \varepsilon$$

$$F \rightarrow iF'$$

$$F' \rightarrow iF' \mid \varepsilon$$

(2) 给出改造后文法的每一个非终结符的 First 集合与 Follow 集合:

$$\text{First}(E) = \{ [\} \quad \text{First}(E') = \{ [, \varepsilon \} \quad \text{First}(F) = \{ i \} \quad \text{First}(F') = \{ i , \varepsilon \}$$

$$\text{Follow}(E) = \{ \# \} \quad \text{Follow}(E') = \{ \# \} \quad \text{Follow}(F) = \{ \} \quad \text{Follow}(F') = \{ \}$$

(3) 给出改造后文法的预测分析表。

	i	[]	#
E		$E \rightarrow [F]E'$		
E'		$E' \rightarrow E$		$E' \rightarrow \varepsilon$
F	$F \rightarrow iF'$			
F'	$F' \rightarrow iF'$		$F' \rightarrow \varepsilon$	

六 设有文法 G (S):

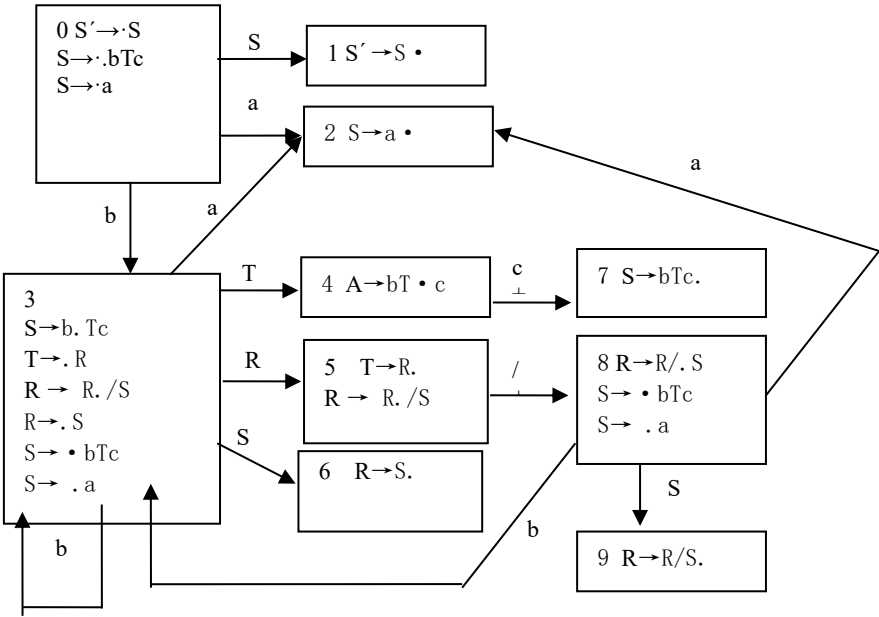
$$S \rightarrow bTc \quad S \rightarrow a \quad T \rightarrow R \quad R \rightarrow R/S \quad R \rightarrow S$$

(1) 构造识别文法所有活前缀的 DFA;

(2) 构造其 SLR 分析表, 判断该文法是否为 SLR (1) 文法。

项目集:

$$(0) S' \rightarrow S \quad (1) S \rightarrow bTc \quad (2) S \rightarrow a \quad (3) T \rightarrow R \quad (4) R \rightarrow R/S \quad (5) R \rightarrow S$$



Follow(S') = {#}
Follow(S) = {c, #}
Follow(T) = {c}
Follow(R) = {c}

QQ: 321556185

状态	ACTION						GOTO		
	a	b	c	/		#	S	T	R
0	S2	S3					1		
1						acc			
2			r2	r2		r2			
3	S2	S3					6	4	5
4			S7						
5			r3	S8					
6			r5	r5					
7			r1	r1		r1			
8	S2	S3					9		
9			r4	r4					

是 SLR（1）文法。

七. 为文法:

$$S \rightarrow aBS \mid bAS$$

$$B \rightarrow aBB \mid b$$

$$A \rightarrow bAA \mid a$$

给出一翻译模式, 统计句子中的 a 的个数和 b 的个数。

$$S \rightarrow aBS \quad S.na := B.na + S1.na + 1$$

$$S.nb := B.nb + S1.nb$$

$$S \rightarrow bAS \quad S.na := A.na + S1.na$$

$$S.nb := A.nb + S1.nb + 1$$

$$B \rightarrow aBB \quad B.na := B1.na + B2.na + 1$$

$$B.nb := B1.nb + B2.nb$$

$$B \rightarrow b \quad B.na := 0$$

$$B.nb := 1$$

$$A \rightarrow bAA \quad A.na := A1.na + A2.na$$

$$A.nb := A1.nb + A2.nb + 1$$

$$A \rightarrow a \quad A.na := 1$$

$$A.nb := 0$$

八. 对以下三地址代码:

I:=1

read J, K

L: M:=K*I

N:=J*I

Q:=M*N

write Q

I:=I+1

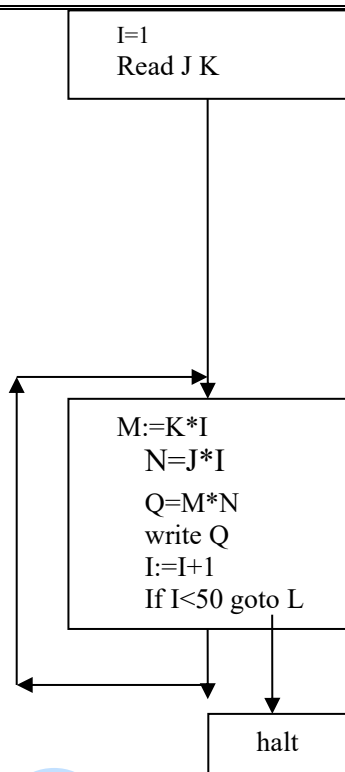
If I<50 goto L

halt

(1) 划分基本块, 并画出程序流程图;

(2) 进行最大可能的循环优化。

程序流图:



QQ: 321556185