

# 编译原理题目解

诗云：太初有道，道曰全真。全真七子，星聚软院<sup>1</sup>。

重阳归天<sup>2</sup>，分道七篇。各执一篇，授业终南。

当是时也，正乃王道长献青真人讲道西电，秘授编译原理心法一篇，以飨众生。

无量寿佛！

今晚 Li Huan

## 1、填空题（30分）

1.1 以阶段划分的编译器中，语法分析 阶段以记号流为输入，语义分析 阶段以语法树为输入。

<sup>1</sup> 此处作者指的是目前西电软件学院七位有德道长：总掌教大护法武道长、首席副掌教顾道长、刘道长讳西洋、沈道长讳沛意、王道长讳献青、高道长讳海昌、陈道婆讳静玉。此乃 huan 说，另有其他版本。不再赘述。

<sup>2</sup> 此处指软件学院开山祖师陈真人讳平。归天意本离世，此处特指升仙为四大护法尊之一。

1.2 有正规式  $P=ab$  和  $Q=cd$  则  $L(QP)=\{cda, cdb\}$  ,  $L((P|Q)Q)=\{acd, bcd, cdcd\}$  。

1.3 有两个因素使得有限自动机是不确定的，一个是 具有  $\epsilon$  状态转移，另一个是 对同一字符，可能有多于一个的下一状态转移。

1.4 词法分析器有四个作用，请给出其中的任意两个：识别记号并交给语法分析器/滤掉源程序中的无用成分/处理与具体平台有关的输入/调用符号表管理器或出错管理器。

1.5 一个定义正确的上下文无关文法，非终结符集合和终结符集合的交集为空，所有出现在产生式左部的文法符号均是 非终结符，仅出现在产生式右部的文法符号均是 终结符。

1.6 编译源程序的过程中，发现函数定义末尾缺少花括号，该情况是 语法 错误；发现除数为 0，该情况是 语义 错误。

1.7 推导  $S \Rightarrow ?H \Rightarrow ?FTP \Rightarrow ?FTc \Rightarrow ?Fbc \Rightarrow ?abc$  是 最右/规范 推导。

1.8 产生式  $F \rightarrow A * F | A$  提取左因子的结果为  $F \rightarrow AF'$   $F' \rightarrow *F | \epsilon$ 。

1.9 对于算术表达式 “ $a*b+c$ ”，当采用预测分析方法时，接受格局中的“当前剩余输入”应该 为空，初始格局中的“当前剩余输入”应该是  $a*b+c$ 。 1.10 最

左归约是最右推导/规范推导的逆过程，每步直接归约均是用产生式左部非终结符替换右句型中的句柄，直到归约为文法开始符号。

1.11 在引用调用的参数传递方式中，调用时传递的是实参的地址，要求实参必须是左值，过程内部对形参的修改等价于对实参的修改。

1.12 假定运算+与\*都是左结合的，且运算\*比运算+优先级高，则算术表达式  $x+y*(u+v)$  的后缀式是 $xyuv+*+$ 。

1.13 拉链-回填技术是语法制导翻译过程中使用的一种基本技术，其基本思想是当三地址码中的转向不确定时将所有转向同一地址的三地址码拉成一个链，而一旦所转向的地址被确定，则为此链上所有的三地址码回填入此地址。

## 2、简答题（20分）

2.1 简述语言的语法和语义，并举一个实际的例子加以说明。

答：语法规定了句子形成的规则，表述了语言的形式，或者说语言的样子和结构，也被称为语法规则。根据语法规则可以识别记号流中的语言结构，也被称为语法分析。语义揭示了语言本身的含义、施加于语言结构上的限制或要执行的动作。

例如“猫吃老鼠”和“老鼠吃猫”都是语法正确的句子，但后者表述的语义不对。

（自己组织语言即可）

2.2 如果一个集合中的元素都是长度不小于 1 且均不以 **ab** 开始的 **a**、**b** 串，请给出描述该集合的正规式。

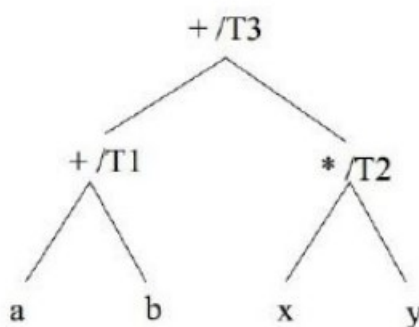
答： $a| (aalb)(alb)^*$

2.3 语法分析器在编译器中应完成什么任务？答：语法分析器根据语法规则识别出记号流中的结构，并构造一棵能够正确反映该结构的语法树。检查输入中的错误，调用出错管理器进行适当处理。

2.4 给定文法  $G: C \rightarrow C h T | T \quad T \rightarrow T a F | F \quad F \rightarrow v$  请给出该文法的终结符集合、非终结符集合，并指出文法的开始符号。

答：终结符： $h、a、v$  非终结符： $C、T、F$  开始符号： $C$

2.5 给出下图中的树对应的三地址码序列。



解：

(+, a, b, T1)
(+, x, y, T2)
(+, T1, T2, T3)

2.6 假设数组下标从 0 开始，对于有 5 行 6 列的数组  $a[5][6]$ ，已知该数组的存储空间首地址为  $a$ ，每个元素占用存储空间大小为  $w$ ，请给出数组以行为主存放时元素  $a[2][3]$  的地址。

答： $a + (2*6+3) * w = 15w + a$

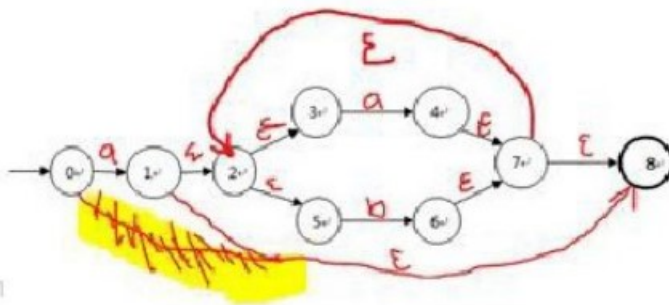
### 3、计算题（50 分）

3.1 给定正规式  $R = a(alb)^*$

<1> 用 Thompson 算法构造识别  $L(R)$  的 NFA  $N$ ；

<2> 用“子集法”把  $N$  确定化（写出完整过程），得到识别  $L(R)$  的 DFA  $D$ ；

<3> 如果  $D$  不是最简 DFA，请找出最简 DFA  $D'$ 。答：<1>笔误：红线一端从 1 开始。



<2>

$E_{\text{闭包}}(\{0\}) = \{0\}$  A

$E_{\text{闭包}}(\text{move}(A,a)) = \{1,2,3,5,8\}$  B

$E_{\text{闭包}}(\text{move}(B,a)) = \{2,3,4,5,7,8\}$  C

$E_{\text{闭包}}(\text{move}(B,b)) = \{2,3,5,6,7,8\}$  D

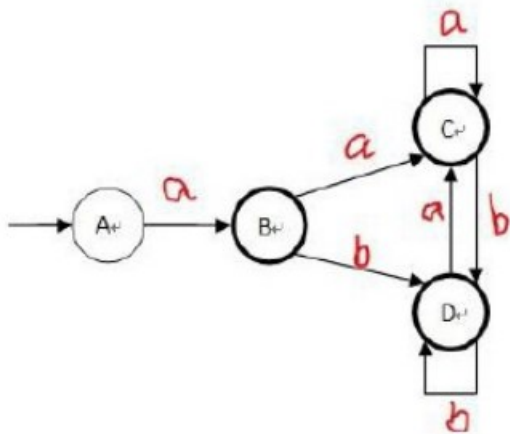
$E_{\text{闭包}}(\text{move}(C,a)) = \{2,3,4,5,7,8\}$  C  $E_{\text{闭包}}(\text{move}(C,b)) = \{2,3,5,6,7,8\}$

$D$   $E_{\text{闭包}}(\text{move}(D,a)) = \{2,3,4,5,7,8\}$  C

$E_{\text{闭包}}(\text{move}(D,b)) = \{2,3,5,6,7,8\}$  D

	a	b
A	B	-
B	C	D
C	C	D
D	C	D

DFA:

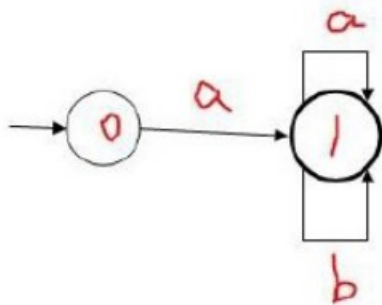


<3>

Move(C,a)=C Move(C,b)=D Move(D,a)=C Move(D,b)=D

B、C、D 不可分。合并为一。

DFA A':



3.2 给定文法 G :

$B \rightarrow B \& C \mid C$

$C \rightarrow E < E \mid E \ E \rightarrow \sim E \mid n$  和

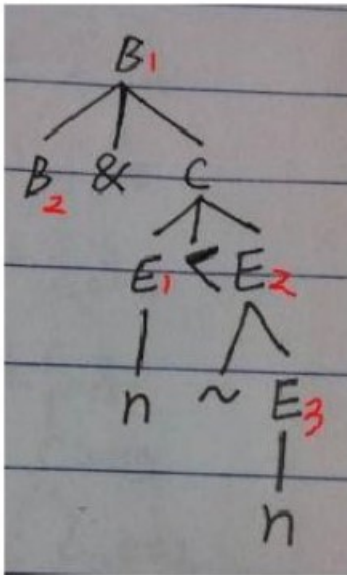
右句型“ $B \& n < \sim n$ ”。

<1>画出该句型对应的分析树； <2>指出句型中的所有短语、直接短语和句柄。



解：

<1>



<2> 短语： $B \& n < \sim n(B1)$ 、 $n < \sim n(C)$ 、 $\sim n(E2)$ 、 $n(E1)$ 、

$n(E3)$  直接短语： $n(E1)$ 、 $n(E3)$  句柄： $n(E1)$

3.3 给定文法  $G$  的拓广文法如下：

$S' \rightarrow S$

$S \rightarrow E \$$

$E \rightarrow id$

$E \rightarrow id ( E )$

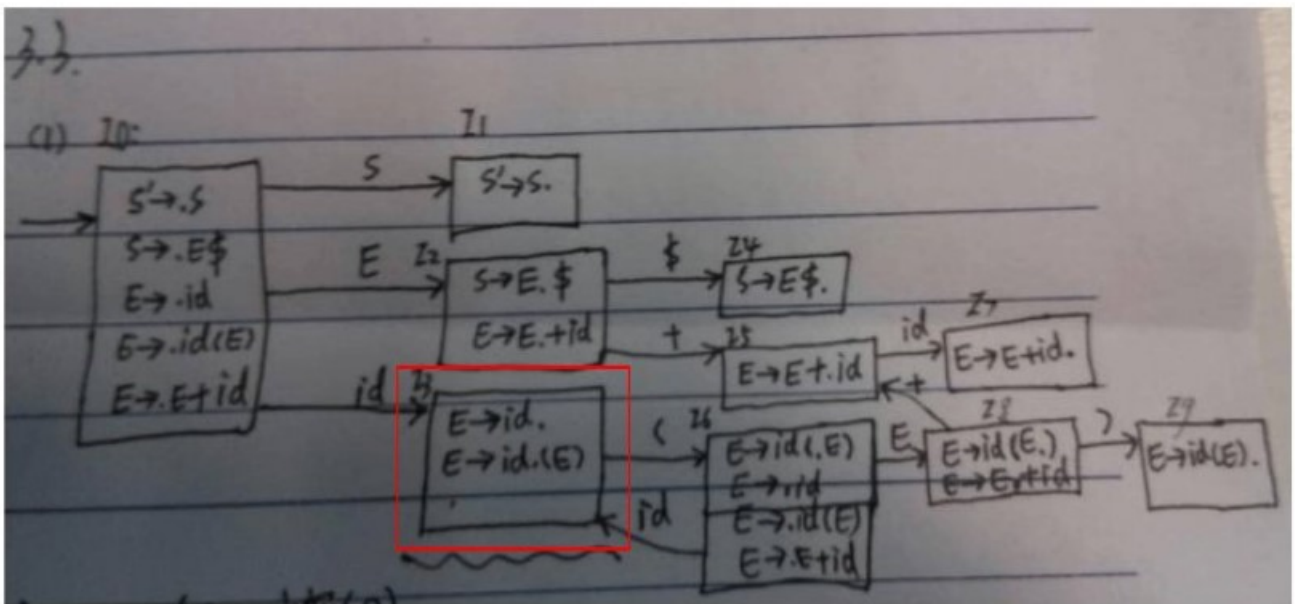
$E \rightarrow E + id$

<1>构造识别  $G$  所有活前缀的 DFA；

<2> $G$  是 SLR(1)文法吗？为什么？

<3> $G$  是 LL(1)文法吗？为什么？若不是，请改写为等价的 LL(1)文法。解：

&lt;1&gt;



&lt;2&gt;是

$I_3$  中存在移进规约冲突。

$FOLLOW(E) \cap ( = \emptyset$ , 则冲突可解决。因此, 是 SLR(1) 文法。

<3> 不是因为存在左递归和左因子。

$E \rightarrow E + id \mid id(E)$  消除左递归 :

$E \rightarrow idE' \mid id(E)E'$      $E' = +idE' \mid \epsilon$  消除左因子 :

$E \rightarrow idF$      $F \rightarrow E' \mid (E)E'$

3.4 给定上下文无关文法和语义规则如下 :

$S \rightarrow aS1a \{ S.count := S1.count + 2; \}$

$\mid bS1b \{ S.count := S1.count + 2; \}$

$\mid C \{ S.count := C.count; \}$

$C \rightarrow cC1 \{ C.count := C1.count + 1; \}$

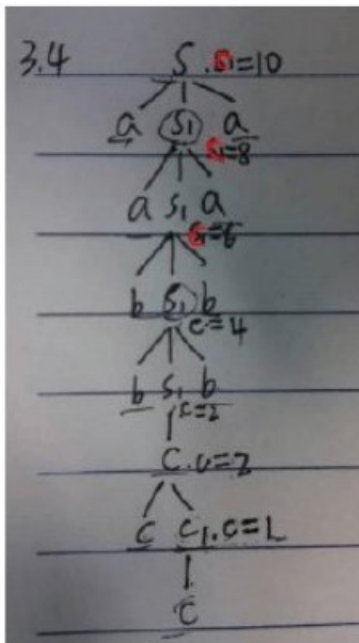


1 c { C.count := 1; }

<1> 画出输入序列 aabbccbbaa 的分析树；

<2> 根据语义规则标注分析树上对应文法符号的.count 值；解：

<1>



注：.c 就是.count，此处为简略表示。

<2>

如上图示。

3.5 忽略过程参数的快排序的部分 Pascal 声明代码如下：

```
program sort; var
```

```
  a:array[10]of integer;
```

```
  x:integer; procedure
```

```
  quicksort;
```

**var i,v:integer; function**

**partition:integer; var**

**i,j:integer;**

<1> 给出上述代码中三个过程（**sort**、**quicksort** 及 **partition**）的嵌套层次；

<2> 给出上述定义对应的嵌套层次的符号表及每个符号表中的符号（假设每个

整型数占用 4 个单元）。解：

