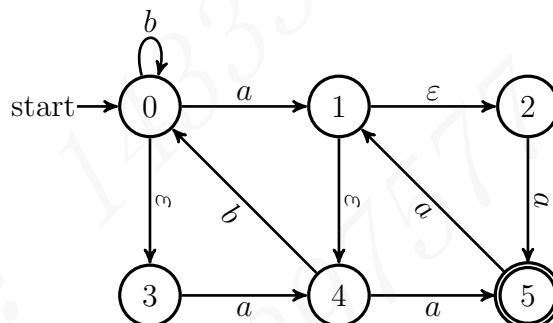


武汉大学计算机学院2009-2010学年第一学期 2007级《编译原理》考试试题

学号：_____ 姓名：_____ 成绩：_____

注意：请将答案全部写在答题纸上，并注明题号！

一、 设NFA N 的状态转换图如下所示： (20分, 5+5+5+5)



- (1) 试写出NFA N 接受字符串“baaabaa”的过程；
- (2) 设用子集构造法求出的与NFA N 等价的DFA M 有三个状态 A 、 B 和 C ，其中 $A = \varepsilon\text{-closure}(\{0\})$, $D\text{trans}(A, a) = B$ ，试求与状态 A 、 B 和 C 所对应的NFA N 的状态集，并画出DFA M 的状态转换图；
- (3) 试设计正规表达式 r ，使得 $L(r) = L(N)$ ；
- (4) 试用自然语言描述 N 所生成的语言。

二、 设文法 $G(S)$ 定义如下： (30分, 每小题5分)

$$\begin{aligned} S &\rightarrow abA \mid b \\ A &\rightarrow Saa \mid baa \end{aligned}$$

- (1) 试写出语句“abababbabaaaaa”的一个最右推导并画出其对应的语法树；
- (2) 试描述文法 $G(S)$ 所生成的语言 $L(G)$ ；
- (3) 试对文法 $G(S)$ 的每个非终结符求First集合和Follow集合；
- (4) 试构造文法 $G(S)$ 的LL(1)分析表，从而说明该文法不是LL(1)文法；
- (5) 设计一个与 $G(S)$ 等价的LL(1)文法；
- (6) 试简要说明文法 $G(S)$ 不是LR(1)文法。

三、 设有简化后的C语言表达式文法 $G(E)$ 定义如下： (10分, 5+5)

$$E \rightarrow E + E \mid E = E \mid *E \mid id$$

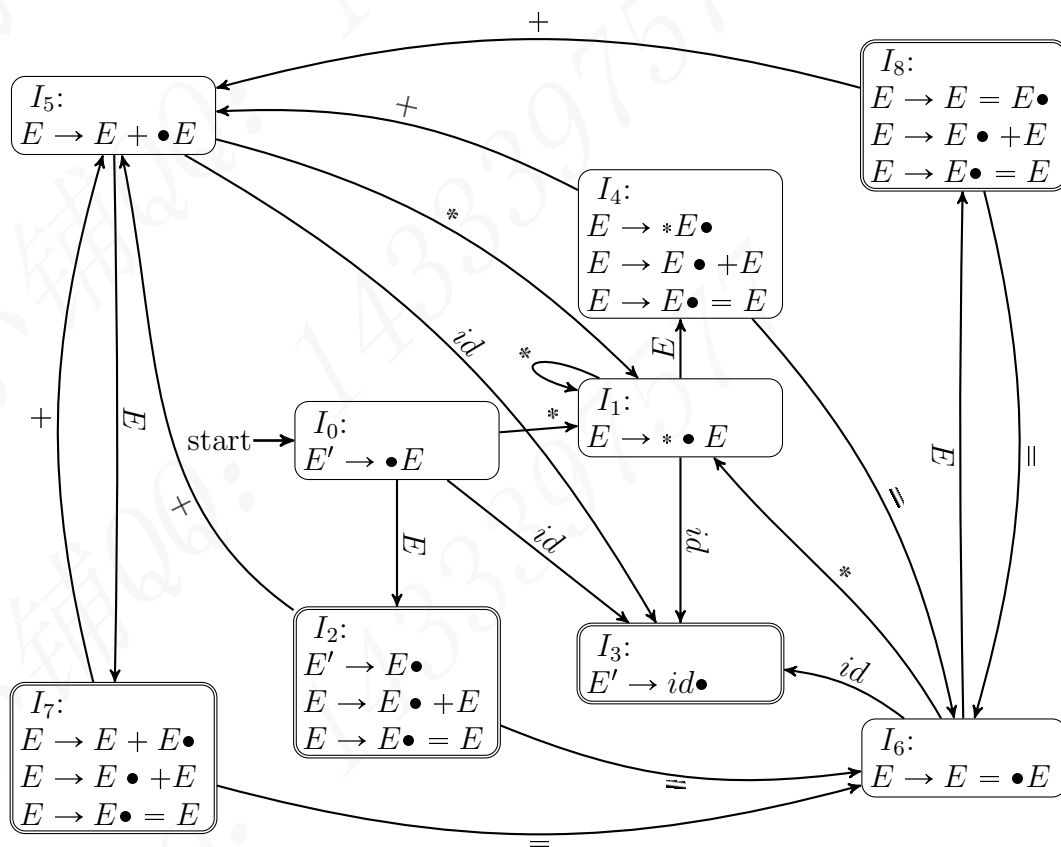
其中：‘ id ’，‘ $=$ ’，‘ $+$ ’和‘ $*$ ’为终结符，“ $*$ ”为一元间接访问运算符；“ $+$ ”为二元加法运算符；“ $=$ ”为二元赋值运算符；

- (1) 试说明该文法是二义文法；
- (2) 设加法运算为左结合，赋值运算为右结合，且间接访问运算优先级高于加法运算，加法运算的优先级高于赋值运算，试设计一个与文法 $G(E)$ 等价的无二义文法，使得其运算的优先级和结合次序与上述规定一致。

四、 设题三的拓广文法 $G(E')$ 定义如下： (20分, 5+5+5+5)

$$\begin{aligned}
 E' &\rightarrow E & (0) \\
 E &\rightarrow E + E & (1) \\
 &| E = E & (2) \\
 &| * E & (3) \\
 &| id & (4)
 \end{aligned}$$

文法 $G(E')$ 的识别活前缀LR(0)项目自动机如下图所示(注意每个状态仅列出了核心项目):



- (1) 试分别求出活前缀“ $E + **$ ”和“ $*E = E$ ”对应的有效项目集；
- (2) 由于该文法是二义文法，因此其SLR分析表一定有移进/归约或归约/归约冲突，请指出哪些状态有怎样的冲突；
- (3) 试画出该文法的SLR分析表，使得其运算的优先级和结合次序与题三(2)的规定一致；

(4) 利用你的分析表写出输入表达式“ $id = *id + id$ ”的分析过程。

五、根据C语言的语义，只有第一个运算量有左值时，赋值运算才有定义。现对题三的表达式文法 $G(E)$ 非终结符 E 设计两个属性 has_lvalue 和 is_ok ，其取值均为布尔常量 $\{true, false\}$ ，并作如下规定：属性 $E.has_lvalue$ 为 $true$ ，当且仅当 E 是 id ，或 E 所代表的表达式的最后一次运算为间接访问运算 $*$ ，如， id ， $*E$ 和 $**E$ 有左值；而 $id = id$ 和 $id + id$ 没有左值。 $E.is_ok$ 的属性为 $false$ ，当且仅当 E 所代表的表达式有一个赋值运算的子表达式，且该子表示的第一运算量没有左值，如当 E 为 $id + id = id$ 或 $id + *id = id$ 时， $E.is_ok = false$ 。(10分，5+5)

(1) 试为 $G(E)$ 设计表达式求值的语法制导定义；

(2) 画出语句 $*id = id + id = id$ 的附注语法树(即在语法树上的每个结点标注其对应的属性值)。

六、设有如下Pascal程序片段：

(5分)

```
repeat
  x := x + 1;
  if (x = 10 and x = 20) then
    x := x + 10
  else
    continue;
  if (x = 100) then break;
until (x < 0 or x > 110);
```

其对应的三地址码如下所示：

16: t0 := x + 1	12: goto l_
x := t0	14: if (x = 100) goto l_
if (x = 10) goto l_	goto l_
goto l_	13: goto l_
10: if (x = 20) goto l_	17: if (x < 0) goto l_
goto l_	goto l_
11: t1 := x + 10	15: if (x > 110) goto l_
x := t1	goto l_
goto l_	18:

试为其中的goto填上正确的标号编号。

注意：第七题见下页！

七、 设有如下C语言程序:

(5分)

```
#include <stdio.h>
void foo(int c)
{
    int * cp = &c;
    *(cp - 2) -= 4;
    printf("a = %d\n", c);
}
int main()
{
    int a = 11;
    int b = 22;
    foo(a);
    printf("a = %d\n", a);
    return 0;
}
```

在Intel x86/Linux下用GCC编译没有任何警告, 运行编译后的程序输出如下结果并正常结束:

```
a = 11
a = 22
```

试简要分析该程序为什么会输出上述结果。