### 计算机学院、网安学院2020－2021学年第一学期

### 本科生编译系统原理期末考试试卷(A卷)

专业：\_\_\_\_\_\_\_\_\_\_\_\_\_\_年级：\_\_\_\_\_\_\_\_\_\_\_\_\_\_学号：\_\_\_\_\_\_\_\_\_\_\_\_\_\_

姓名：\_\_\_\_\_\_\_\_\_\_\_\_\_\_成绩：\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |
| --- |
| **得 分** |
|  |

1. 单项选择题（每空2分，共24分）
   * + 1. C++编译器进行常数的类型转换是在 C 阶段，过滤注释是在  
           A 阶段，消除公共子表达式是在 D 阶段。

A．词法分析 B．语法分析

C．语义分析 D．代码优化

* + - 1. GCC-ARM将C++程序转换为ARM机器码程序，模拟器在X86平台上模拟执行ARM机器码程序是采用读取一条指令——转换为相应的X86机器码——执行，如此往复的方式，则GCC-ARM是一种\_\_\_\_B\_\_\_\_\_，模拟器是一种\_\_\_\_D\_\_\_\_\_。

A．预处理器 B．编译器

C．链接器 D．解释器

* + - 1. 在词法分析和语法分析中，****为\_\_\_D\_\_\_\_\_\_。

A．终结符

B．非终结符

C．空符号

D．空符号串

* + - 1. 正则表达式间的“**|**”运算支持交换律，这是因为正则表达式本质上描述的是\_\_\_\_D\_\_\_\_\_\_\_。

A．符号 B．符号集合

C．符号串 D．符号串集合

* + - 1. 下列语言可以用DFA识别的是\_\_\_\_D\_\_\_\_\_，可用CFG识别的是\_\_\_\_C\_\_\_\_\_。

A．形如**xx**的0、1串集合

B．形如**anbncn**（n≥1）的串的集合

C．正则表达式集合

D．SLR(1)文法的活前缀集合

* + - 1. 对下面CFG，说法**错误的**是\_\_\_\_D\_\_\_\_\_。

**S → 0X X → Y1 Y → Y0 | 0 Z → X1**

A．**Z**是无用的 B．与**00+1**对应相同的语言

C．符合算符文法定义 D．**001**是其活前缀

* + - 1. 如果将物种视为类型，且对类型采用**名字等价**判定，则“披着羊皮的灰太狼”会被认为\_\_\_\_A\_\_\_\_。

A．是类型“羊”

B．是类型“狼”

C．是类型“灰太狼”

D．可转换为类型“红太狼”

* + - 1. L-属性定义的翻译更容易和\_\_\_\_A\_\_\_\_\_相结合。

A．预测分析 B．算符优先分析

C．SLR分析 D．规范LR分析

|  |
| --- |
| **得 分** |
|  |

1. 设计题（每题6分，共24分）

1. 描述下面正则表达式接受什么符号串集合（注意：转义符\表示后面的**"**应视为普通字符，其他符号均为正则表达式运算符）。

**\"([^\"])\*\"**

答：接受双引号包围的字符串，字符串中不包含双引号。

2. 设计接受IPv4地址的正则表达式（可用正则定义）。

答：

D**→**[0-9]

T**→[**1-9][0-9]

H**→**1[0-9][0-9]

B**→**2([0-4][0-9] | 5[0-5])

S**→**D | T | H | B

IPAddr**→**S\.S\.S\.S

3. 设计接受语言{a*i*b*j*a*k*b*l* | *i*+*j*=*k+l*, *i*, *j*, *k*, *l*>=0}的上下文无关文法。

答：

S→aSb | A | B | M

A→aAa | M

B→bBb | M

M→bMa | 

4. 设计接受C++数组声明语句的上下文无关文法，其中数组元素类型限定为int、char及它们的指针，数组维数可以是任意维。

答：

D→T id M ;

T→int | char | T\*

M→M [num] | [num]

|  |
| --- |
| **得 分** |
|  |

1. （22分）对下面的正则表达式。

**(0 | 1)\*110(0 | 1)\***

1. 用**Thompson构造法**将其转换为NFA，识别0111010。（8分）



识别0111010过程：  
0🡪1🡪2🡪3🡪6🡪1🡪4🡪5🡪6🡪7🡪8🡪9🡪10🡪11🡪14🡪15🡪16🡪11

🡪12🡪13🡪16🡪17

2. 用**子集构造法**将得到的NFA转换为DFA，画出最终的状态转换图，识别0111010。（10分）

\_closure({0})={0, 1, 2, 4, 7}=A

A, 0\_closure(({0, 1, 2, 4, 7}, 0))={1, 2, 3, 4, 6, 7}=B

A, 1\_closure(({0, 1, 2, 4, 7}, 1))={1, 2, 4, 5, 6, 7, 8}=C

B, 0\_closure(({1, 2, 3, 4, 6, 7}, 0))={1, 2, 3, 4, 6, 7}=B

B, 1\_closure(({1, 2, 3, 4, 6, 7}, 1))={1, 2, 4, 5, 6, 7, 8}=C

C, 0\_closure(({1, 2, 4, 5, 6, 7, 8}, 0))={1, 2, 3, 4, 6, 7}=B

C, 1\_closure(({1, 2, 4, 5, 6, 7, 8}, 1))={1, 2, 4, 5, 6, 7, 8, 9}=D

D, 0\_closure(({1, 2, 4, 5, 6, 7, 8, 9}, 0))={1, 2, 3, 4, 6, 7, 10, 11, 12, 14, 17}=E

D, 1\_closure(({1, 2, 4, 5, 6, 7, 8, 9}, 1))={1, 2, 4, 5, 6, 7, 8, 9}=D

E, 0\_closure(({1, 2, 3, 4, 6, 7, 10, 11, 12, 14, 17}, 0))={1, 2, 3, 4, 6, 7, 11, 12, 13, 14, 16, 17}=F

E, 1\_closure(({1, 2, 3, 4, 6, 7, 10, 11, 12, 14, 17}, 1))={ 1, 2, 4, 5, 6, 7, 8, 11, 12, 14, 15, 16, 17}=G

F, 0\_closure(({1, 2, 3, 4, 6, 7, 11, 12, 13, 14, 16, 17}, 0))=F

F, 1\_closure(({1, 2, 3, 4, 6, 7, 11, 12, 13, 14, 16, 17}, 1))=G

G, 0\_closure(({1, 2, 4, 5, 6, 7, 8, 11, 12, 14, 15, 16, 17}, 0))=F

G, 1\_closure(({1, 2, 4, 5, 6, 7, 8, 11, 12, 14, 15, 16, 17}, 1))= { 1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15, 16, 17}=H

H, 0\_closure(({1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15, 16, 17}, 0))={1, 2, 3, 4, 6, 7, 10, 11, 12, 13, 14, 16, 17}=I

H, 1\_closure(({1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15, 16, 17}, 1))=H

I, 0\_closure(({1, 2, 3, 4, 6, 7, 10, 11, 12, 13, 14, 16, 17}, 0))=F

I, 1\_closure(({1, 2, 3, 4, 6, 7, 10, 11, 12, 13, 14, 16, 17}, 0))=G



识别0111010过程：A🡪B🡪C🡪D🡪D🡪E🡪G🡪F

3. 将DFA最小化，画出最终的状态转换图。（4分）

初始化{A, B, C, D}，{E, F, G, H, I}，后者不可再分

0将前者分裂为{A, B, C}和{D}，1将前者分裂为{A, B}和{C}，至此不可再分



|  |
| --- |
| **得 分** |
|  |

1. （15分）对下面文法：

**S→Ac | Bc**

**A→Aa | **

**B→Bb | **

1. 指出其终结符集合、非终结符集合、开始符号（3分）

答：

终结符集合：a、b、c

非终结符集合：S、A、B

开始符号：S

2. 消除文法左递归（4分）

S→Ac | Bc

A→aA | 

B→bB | 

3. 构造预测分析表，对句子aac进行分析（8分）。

答：

FIRST(S)={a, b, c} FOLLOW(S)={$}

FIRST(A)={a, } FOLLOW(A)={c}

FIRST(B)={b,  FOLLOW(B)={c}

预测分析表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | a | b | c | $ |
| S | S→Ac | S→Bc | S→Ac  S→Bc |  |
| A | A→aA |  | A→ |  |
| B |  | B→bB | B→ |  |

分析过程：

|  |  |  |
| --- | --- | --- |
| 栈 | 输入缓冲 | 动作 |
| $S | aac$ | S→Ac |
| $cA | aac$ | A→aA |
| $cAa | aac$ |  |
| $cA | ac$ | A→aA |
| $cAa | ac$ |  |
| $cA | c$ | A→ |
| $c | c$ |  |
| $ | $ | accept |

|  |
| --- |
| **得 分** |
|  |

1. （5分）对下面流图，指出所有回边及每条回边对应的循环包含哪些顶点。



答：

回边2🡪2，对应循环包括2

回边3🡪3，对应循环包括3

回边5🡪2，对应循环包括2、3、4、5

|  |
| --- |
| **得 分** |
|  |

1. （10分）下面文法描述了类型表达式，设计语法制导定义实现构造类型表达式对应的表达式树。使用属性p保存每个语法符号对应的表达式树的根节点指针，假设已有辅助函数mkleaf(basic\_type)（及mkleaf(num, val)）和mknode(op, child1, child2)分别为基本类型（及立即数）和构造类型创建叶节点和内部节点，直接使用即可。

T → integer | char | real | void | array(num, T) | T '╳' T | pointer(T) | T '→' T

答：

T → integer { T.p = mkleaf(integer); }

T → char { T.p = mkleaf(char); }

T → real { T.p = mkleaf(real); }

T → void { T.p = mkleaf(void); }

T → array(num, T1) { T.p = mknode(array, mkleaf(num, num.val), T1.p); }

T → T­1 '╳' T2 { T.p = mknode(prod, T1.p, T2.p); }

T → pointer(T1) { T.p = mknode(pointer, T1.p, NULL); }

T → T­1'→' T2{ T.p = mknode(func, T1.p, T2.p); }