### 计算机学院、网安学院2023－2024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）

（空格末尾的编号与答题卡中编号对应）

* + - 1. LLVM编译器生成IR代码是在 ⑴阶段。

A．词法分析 B．语义分析

C．中间代码生成 D．代码优化

* + - 1. LLVM编译器发现函数调用的实参与函数定义的形参不匹配是在  
          ⑵阶段。

A．词法分析 B．语义分析

C．中间代码生成 D．代码优化

* + - 1. LLVM编译器消除无用代码是在 ⑶阶段。

A．词法分析 B．语义分析

C．中间代码生成 D．代码优化

* + - 1. 将可重定位机器码放置到内存并调整其中地址是 ⑷的工作。

A．预处理器 B．编译器

C．链接器 D．加载器

* + - 1. 解决程序中的外部引用是 ⑸的工作。

A．预处理器 B．编译器

C．链接器 D．加载器

* + - 1. 用Thompson算法将一个有a个符号（包括****）、b个运算符的正则表达式r转换为NFA N，则下面说法**正确**的是 ⑹。

A．N中每个状态只可能有一条****边

B．N中每个状态不会有标记字母表中相同符号的多条边

C．N可能有多个终态 D．N最多有(a+b)个状态

* + - 1. 下面可用上下文无关文法描述的语言是\_\_\_\_\_\_\_\_\_\_\_⑺。

A．**{w2w | w是任意0、1串}** B．**{ 0n1n | n≥0 }**

C．**{ 0n1m2n3m | n≥1且m≥1 }** D．以上皆错

* + - 1. 对LR(1)项目[**A→XYZ•, a**]，说法**错误的**是 ⑻。

A．它表示在某个最右句型中**a**跟随**A**

B．它表示在某个最右句型中**a**跟随**XYZ**

C．它表示可以进行归约了，输入缓冲区首符号可以是任意的

D．它表示可以进行归约了，输入缓冲区首符号可以是**a**

* + - 1. 对下面CFG，后面给出的推导是 ⑼。

**S → TU T → 0T1 |  U→1U0 | **

**S⇒TU⇒T1U0⇒T10⇒0T110⇒00T1110⇒001110**

A．最左推导 B．最右推导

C．以上皆错

* + - 1. 对上题CFG，下面推导是 ⑽。

**S⇒TU⇒T1U0⇒0T11U0⇒0T110⇒00T1110⇒001110**

A．最左推导 B．最右推导

C．以上皆错

* + - 1. 在进行代码优化时，应重点考虑哪种程序结构？ ⑾。

A．变量声明 B．赋值

C．条件分支 D．循环

* + - 1. 在采用名字等价判定的高级语言程序中，如果披着羊皮的灰太狼（对原始类型“狼”定义了类型别名“羊”）被认为是“羊”，是否会与喜羊羊（原始类型就是“羊”）无法区分呢？ ⑿。

A．转用结构等价判定来区分两者

B．同一作用域中只允许有一个名为“羊”的类型，后定义者非法

C．通过它们的二进制编码表示来进行区分

D．禁止邪恶的灰太狼，保留善良的喜羊羊

|  |
| --- |
| **得 分** |
|  |

1. 设计题（每题5分，共20分）

1. 设计正则表达式，与下面DFA接受完全相同的符号串集合。



2. 设计正则表达式，接受所有十进制定点数，无多余的前导零和末尾零。即，0.0、123.01、123005.0这样的串是合法的，但00.0、001.000、002345.1000是非法的。

3. 证明下面文法生成的所有二进制串都能被3整除。

***num*** **→ 11 | 1001 | *num* 0 | *num num***

4. 设计CFG，接受与第1小题DFA完全相同的符号串集合。

|  |
| --- |
| **得 分** |
|  |

1. 词法分析算法（24分）

对下面的正则表达式：

**(a\*bc) | (ab\*c) | (abc\*)**

1. 用**Thompson构造法**将其转换为NFA（NFA状态从0开始编号），识别abbc。（6分）

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abbc。（10分）

3. 将DFA最小化，画出最终的状态转换图。（8分）

|  |
| --- |
| **得 分** |
|  |

1. 语法分析算法（16分）

对下面文法：

**S → A a | B b A → d A | d B→ c B | c**

1. 指出其终结符集合、非终结符集合、开始符号（3分）

2. 计算所有非终结符的First函数和Follow函数结果（3分）

3. 构造SLR分析表（从拓广产生式开始、所有产生式从0开始编号），对句子cccb进行分析，画出语法树（10分）。

|  |
| --- |
| **得 分** |
|  |

1. 语法制导翻译（10分）

写出描述类型表达式（支持基本类型char、int、real、bool以及类型构造符array、×、record、pointer和🡪）的CFG，为其设计语法制导定义，构造对应的抽象语法树（叶节点为基本类型、非叶节点为类型构造符，假设已有创建叶节点的辅助函数mkleaf和创建非叶节点的辅助函数mknode，在语法制导定义中直接调用即可）。

|  |
| --- |
| **得 分** |
|  |

1. 中间代码生成（6分）

采用临时变量重用算法将下面表达式翻译为三地址码，临时变量名为t1、t2、…。

**a+a\*(a+(a+a)\*(a+a+a) +a)**