

- 第2页，共8页

得 分

二、 设计题（每题 6 分，共 24 分）

1. GB2312 编码将字符分为 94 个区（编号 1-94），汉字占据 16-87 区。每区 10 行 10 列，按行主次序编号 0-99，其中 0 和 95-99 是六个空位。每个字符采用两字节编码，高字节为区号，低字节为区内位号，区号和位号（先转换为十六进制）再分别加上 0xA0。例如，“王”在 45 区 8 行 5 列，因此其编码为 0xCDF5。设计正则表达式描述 GB2312 中汉字的编码。

2. 英文单词不能以 I、U、V 或 J 结尾，例外是允许三个很古老的英语单词 I、YOU 和 THOU。设计正则表达式描述满足这一规则的字母串（大写）。

3. 设计非二义性上下文无关文法描述布尔表达式，其中基本布尔表达式为常量 **true**、**false** 和变量 **id**，布尔运算有 **and**、**or**、**not** 和(、)。

4. 设计上下文无关文法，生成正则表达式 $0(0|1)^*0$ 所描述的语言。

得 分

三、（22 分）对下面的正则表达式。

$(a | b^*)^*abb$

1. 用 Thompson 构造法将其转换为 NFA。（7 分）

2. 用子集构造法将得到的 NFA 转换为 DFA，写出识别 ababbab 的状态转换序列和识别结果。（9 分）

3. 将 DFA 最小化 (6 分)

得 分

四、 (14 分) 对下面 if 语句的文法 (**i**——if、**e**——else、**a**——代表其他语句的终结符, 忽略了表达式):

(1) 构造 SLR 分析表 (7 分)

(2) 消解分析表中冲突, 对 **ii a e i a e a** 进行语法分析 (7 分)。

$S \rightarrow iS | iSeS | a$

得 分

五、（16 分）

（1）设计上下文无关文法描述类型表达式（注意，不是描述类型），基本类型为 **integer**、**char**、**real**，类型构造符为 **array**、**pointer**、**→**（函数）和 **X**（笛卡儿积）。（6 分）

（2）设计语法制导定义实现类型表达式转换为紧凑二进制编码表示，**integer**、**char**、**real** 编码分别为 **01**、**10**、**11**，**array**、**pointer**、**→**（函数）编码分别为 **01**、**10**、**11**。对类型表达式 **array(num, pointer(char)→integer)**（数组的索引集合用一个常数表示），画出其语法树，利用你设计的语法制导定义对语法树中节点计算属性值，给出最终的翻译结果（10 分）

