

1. 一个典型的编译程序包括：词法分析、\_\_\_\_\_、\_\_\_\_\_、中间代码生成、机器无关代码优化、\_\_\_\_\_、机器相关代码优化；其中“三地址代码生成”属于\_\_\_\_\_。
2. 已知正则表达式  $a^*(bb)^*b$ ，请给出其表示的语言\_\_\_\_\_。
3. 已知文法： $E \rightarrow E+T \mid E-E \mid T, T \rightarrow T * F \mid T / F \mid T, F \rightarrow (E) \mid id$ , 请给出  $id / (id-id)$  的最左推倒\_\_\_\_\_和语法树\_\_\_\_\_；给出  $T / (F-id)$  的句柄\_\_\_\_\_以及对该句柄规约所使用的产生式。
4. 语言  $\{a^n cb^m \mid n, m \geq 1\}$  \_\_\_\_\_用正则表达式描述，\_\_\_\_\_上下文无关文法描述（“可以”或“不可以”）
5. 对于一个按行存放的整型数组  $B[i][j][k][l]$ ，下表  $i$  的取值范围是 0-5,  $j$  的取值范围是 2-4,  $k$  的取值范围是 1-4,  $l$  的取值范围是 4-8. 假设每个元素占 8 个字节，数组  $B$  从编号 1000 的字节开始存放。那么元素  $B[4][2][3][7]$  的起始位置为第\_\_\_\_\_字节。
6. 为文法  $S \rightarrow SA \mid A, A \rightarrow I := E, E \rightarrow E+F \mid F, F \rightarrow (E) \mid i$  设计预测分析表， $FIRST(A) = \{ \_\_\_\_\_\_ \}$ ;  $FIRST(E) = \{ \_\_\_\_\_\_ \}$ ;  $FOLLOW(S) = \{ \_\_\_\_\_\_ \}$ ;  $FOLLOW(E) = \{ \_\_\_\_\_\_ \}$ 。
7. 在全局机器无关的优化中有哪些常用的手段\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。在数据流分析中有哪些基于数据流方程的具体分析技术\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
8. 引入计数垃圾回收的基本思路是：\_\_\_\_\_；这种方法有什么不足：\_\_\_\_\_。
9. 按照课本介绍的方法，为正则表达式  $(ba^?)^*$  构造对应的 NFA\_\_\_\_\_；并使用子集构造法将该 NFA 转换为对应的 DFA\_\_\_\_\_。（给出图和转换表，最小化）
10. 已知文法： $S \rightarrow aA \mid Ab, A \rightarrow Sc \mid d$ , 此文法是否存在左递归？若存在，请消除其左递归（给出结果文法）；否则，说明理由。
11. 原本无冲突的 LR(1) 分析表在合并同心集后得到的 LALR(1) 分析表，新的表中不可能存在哪种冲突，为什么。
12. 已知文法  $E \rightarrow E+T \mid T, T \rightarrow TF \mid F, F \rightarrow F * (E) \mid a \mid b$ , 给出 LR(1) 项集  $[F \rightarrow (\cdot E), +]$  的闭包；给出上述项集闭包相对于文法符号  $E$  的后继 LR(1) 项集闭包。
13. 已知文法  $E \rightarrow E-T \mid T, T \rightarrow num \mid num., num$  请给出一个 SDD，来确定减法表达式的类型；注： $T$  有综合属性  $type$ （属性值分为  $integer$  和  $float$  两种）， $E$  有综合属性  $type$ 。可以使用函数  $getType(type\_1, type\_2)$  来为减法表达式获取类型。
14. 请给出 `repeat` 语句的中间代码翻译方案。已知其语法如下： $S \rightarrow repeat S1 while B$  假设其他类型的语句和表达式的翻译方案已经生成。注： $S$  具有综合属性  $nextlist$  用来表示跳转指令列表， $B$  具有综合属性  $truelist$  和  $falselist$  来管理布尔表达式跳转代码中的标号，所插入的额外标记非终结符号具有属性  $instr$  表示回填的目标标号。必须使用回填。
15. 给出 DAG 图，标出被 killed 的节点。 $a=b+c, d[i]=e, a=e*g, f=a+e$ 。

16. 已知三地址代码  $b=1, c=c+b, n=10, \text{if } c < n \text{ goto}(6), d=1+n, a=a+b, c=2, \text{if } a < n \text{ goto}(4), t1=4*a, c=a-b, t2=1+n$ , 给出相应的流图（使用三地址代码前的序号来表示对应的三地址代码），并指出其中所有的循环。指出全局公共子表达式，并指出其所在的三地址代码的序号；指出每个循环中的可进行强度消减的归纳变量，并说明如何强度消减。
17. 将三地址代码组成的基本块翻译成机器代码，并给出每条机器代码执行后的寄存器描述符和地址描述符（目标指令从 100 开始编号；假设有三个寄存器可用 R1, R2, R3，使用 getReg() 分配寄存器；其中 a、b、c、d 全局变量；t、u 是临时变量，不需要保存）  
 $t=a*b, u=a-c, d=t+u$  附机器代码格式：取数指令：LD R, var //将内存空间变量 var 的值存入寄存器 R 中。存数指令：ST var, R //将寄存器 R 中的值存入变量 var 的内存空间中。双目运算指令：OP dst, src1, src2 //使用 OP 对寄存器 src1, src2 中的值进行计算，结果存入寄存器 dst 中（OP 包括 ADD, SUB, MUL, DIV）
18. 构造一个正则表达式及与之等价的一个最小化 DFA。它们描述能够被 4 整除的长度不小于 2 的二进制串。假设串的集合为 L，则如 100#L, 000000#L, 0011100#L, 1001! #L。
19. 构造预测分析表，并说明该文法是否是 LL(1) 文法（\$ 为结束符号；必须先写出 FIRST 和 FOLLOW 集合再构造预测分析表）  
 $S \rightarrow uBDz, B \rightarrow Br | w, D \rightarrow EF, E \rightarrow y | \#, F \rightarrow x | \#$
20. 考虑如下的上下文无关文法 CFG，S 为开始符号： $S \rightarrow NV, N \rightarrow a | b | N \& V, V \rightarrow 0 | 1N$ ，若非终结符 N 所表示的长度小于 5（即 0-4）的串有哪些？非终结符 V 所表示的长度小于 5（即 0-4）的串有哪些？考虑串  $b \& 01a \& 0$ ，是否属于此文法产生语言？如果是，给出语法分析树
21. 表达式  $3*5+4$  的后缀表示是  $35*4+$ ，基于文法 G:  $L \rightarrow ES, E \rightarrow E+T | T, T \rightarrow T * F | F, F \rightarrow (E) | \text{digit}$  实现一个可以将表达式转换成后缀表达的不带左递归的翻译方案，并画出表达式  $35*4+$  对应的注释分析树（假设 digit, value 包含其带有的值）
22. 分析流图，给出到达位置 4 变量 a 的定值点有\_\_\_\_\_，到达位置 8 变量 a 的定值点有\_\_\_\_\_，到达位置 6 变量 b 的定值点有\_\_\_\_\_。（图中符号 <- 表示赋值操作，bgt 和 blt 表示比较大小后跳转，答案写出定值点的位置）。