```
源语言 源程序 | 目标语言 目标程序 | 宿主机 宿主语言
              第一章 编译引论
                                分类: 汇编程序 编译程序 解释程序 | 解释执行 编译执行
                               编译过程: 词法分析 语法分析 语义分析与中间代码生成 代码优化 目标代码生成
                               编译程序构造:源语言目标语言与目标机编译方法与工具
                                          巴克斯-诺尔 范式
                                           文法: 非空有限终结符号集 非空有限非终结符号集 开始符号 产生式集
                                          最右推导 = 规范推导
                                          文法等价:产生相同语言的文法等价
                                          文法二义性:一个文法的语言可以对应两个不同的语法分析树(有两个不同的最左/规范推导)
                                                    0型文法(短语结构文法):所有文法都是0型文法 图灵机识别
                                                    1型文法(上下文有关文法):线性有界自动机识别
                                          文法类型
                                                    2型文法(上下文无关文法): 非确定下推自动机识别
                                                    3型文法(正则文法/线性文法):确定有限自动机识别
                                                    数字越大,限制越多,范围越小,越被包含
              第二章 形式语言与自动机理论
                                                     确定有限自动机DFA: 与3型文法(正则文法)——对应 | 状态集 输入字符集 状态转换函数 唯一初态 终态集
                                                     非确定有限自动机NFA:允许不识别字符就转换状态允许转换状态不唯一
                                                     对于任何一个非确定有限自动机,都存在一个确定有限自动机与之对应相等
                                          有限自动机
                                                      非确定有限自动机确定化算法 (子集法)
                                                      确定有限自动机化简:消除无关状态(从开始符号不可能到达的状态)消除等价状态(划分法)
                                         闭包 (closure)
                                                   正规式表示的语言 = 正规集 = 有限自动机识别的语言
                                          正规式
                                                   有限自动机与正规式转换算法
                                单词: 关键字 常数 标识符 运算符 界限符 (属性字分类)
              第三章 词法分析
                                对半互补缓冲区:解决 超前搜索 回退字符 问题
                                状态转换图 | 状态转换表(数据中心法)
                                       自上而下: 推导 | 自下而上: 归约(推导归约互为逆过程)
                                                                   消除左递归(直接|间接)消除递归算法
                                        自上而下不确定的原因: 左递归问题 回溯问题
                                                                            FIRST 集
                                                                   消除回溯
                                                                            同一个产生式的多个候选式 FIRAT 集相交 – 提取左公因子
                                        递归下降分析器
              第四章 语法分析 自上而下
                                               组成:输入串 总控程序 分析栈
                                               FOLLOW 集(FIRST 集中有空时需要考虑)
                                       LL(1)
                                                LL(1)分析表的构造
                                               LL(1)文法: 非二义文法 | 不能有左递归 | 不能有回溯(同一产生式的 FIRST 集无交集)
                                      目前 最常用 最有效
                                       短语 直接短语 句柄(最左直接短语) 通过画语法分析树进行寻找
                                                   优先关系 < = > 只有在语言中相邻的两个运算符间才存在优先关系
                                                   算符优先文法: 文法中没有两个非终结符号相连的产生式
                                                   FIRSTVT LASTVT
                                       算符优先分析
                                                  素短语: 至少有一个终结符号 其中不再包含素短语 | 最左素短语(规范归约的可规约串)
                                                                              优先函数构造算法(分组 连线 标号)
                                                  优先函数表: 优先矩阵的线性化 | 不一定存在
                                                组成: 总控程序(LR分析表) 输入串 分析栈
编译原理
                                                动作: 移进 归约 出错 接受
                                                           活前缀: 规范句型的一个不含句柄之后任何符号的前缀
              第五章 语法分析自下而上
                                                           LR(0)项目: 归约项目 接受项目 移进项目 待约项目
                                                 LR(0)分析
                                                           LR(0)项目集规范簇:识别文法可归前缀 DFA 项目集的全体
                                                                                             DFA 中不能含有 移进-归约 和 归约-归约冲突
                                                           构造 LR(0)分析表
                                       LR 分析
                                                           LR(0) 的 DFA 中含有 移进-归约 和 归约-归约 冲突
                                                                         识别字母等于移进项-移进
                                                SLR(1)分析
                                                           改进: FOLLOW 集
                                                                         识别字母等于归约项 FOLLOW 集-归约
                                                                         否则报错
                                                 LR(1)分析 (规范 LR 分析)
                                                                                                      更新了 LR(0) 的项目-增加了搜索符号即产生式左边非终结符号的 FOLLOW集
                                                                    LR(1)项目集规范簇:识别文法可归前缀 DFA 项目集的全体
                                                            同心项目集: LR(1) 的项目中项目相同仅搜索符不同的项目
                                                 LALR(1)分析
                                                            合并同心项目集:可能产生归约-归约冲突
                                           编译阶段最实质性的工作 对源程序的语义作出解释 引起源程序发生质的变化
                                           包含: 语法树的构造 符号表的组织 类型检查
                                           中间语言:逆波兰表达式(运算符的顺序就是实际计算的顺序 | 所以不需要括号方便计算机计算)
                                                        三元式
              第六章 语义分析与中间代码生成
                                                        间接三元式 (用 间接码表 给出三元式的执行顺序)
                                            N 元式表示法
                                                        四元式(用单独的一项表示每一步计算结果T)
                                                       BZ:条件跳转 BR:无条件跳转
                                           语句翻译-语句标号 AND 拉链反填技术 (标号名 定义与否 地址a)
                                概念:通过重排 删除 合并 改变程序等手段使程序变得高效(时间***空间)
                               优化效果是编译器的一个重要的性能指标
                                             常量合并与传播
                                             公共子表达式的删除
                                             无用赋值的删除
                                             死代码的删除
                                优化技术的分类
                                             无用转移语句的删除
                                             循环不变量或不变代码外提
                                             函数内嵌
                                             循环转换
                                             其他
                                                  基本块:程序中一组顺序无转移执行的语句序列出入口唯一
                                                   入口语句: 1、程序的第一个语句 2、转移语句转移到的语句 3、条件语句后的语句
                                          基本块
                                                  基本块划分:一个入口 到 下一个入口语句(不含)/下一转移语句(含)/程序终止语句(含)
              第八章 代码优化
                                                   未包含在基本块中的语句都是程序不可达语句-直接删除
                                                    控制流图(流图):按照程序的执行过程用有向边把基本块链接起来
                                          控制流图
                                                    流图是一个首节点唯一的有向图
                                                           DAG 表示基本块之间的关系
                                          DAG (无环路有向图)
                                局部优化
                                          必经节点 必经节点集 回边 循环
                                         IN(B): 到达基本块 B 入口时,各个变量的所有定值点集合
                                         GEN(B):在基本块 B 中产生的定值点
                                         KILL(B): 在基本块 B 中重定义的定值点
                                         OUT(B): 到达基本块 B 出口时,各个变量的所有定值点集
                                          ud 链:引用-定值链:能到达 P(引用了变量 A)的 A 的定值点的全体
                                          du 链:定值-引用链:P定值的变量A能到达的引用点的全体
```

编译程序:把用某一种程序设计语言写的源程序翻译成等价的另一种语言程序的程序