

编译原理课程笔记

卢雨轩 19071125

2022 年 2 月 25 日

第一章 绪论

主要内容

- 编译原理及其设计概述 3 学时
- 语言与文法 5 学时
- 词法问题 6 学时
- 语法分析 13 学时 自顶向下 (LL); 自底向上 (LR);
- 语义分析 10 学时
- 运行环境 3 学时 过程调用, 符号管理
- 优化与代码生成 2 学时
- 总结

一、 计算机语言的发展

- 机器语言 Machine Language
 - 二进制代码 Binary Code
- 汇编语言 Assemble Language
 - 二进制代码与助记符
 - 接近计算机硬件指令系统
- 高级语言 High Level Language
 - 语句定义数据、描述算法
- 命令语言 Command
 - 功能封装
- 高级语言的分类
 - 命令式语言 Imperative Language
 - * Fortran, Basic, Pascal, C, COBOL, ALGOL
 - 函数式语言 Functional Language
 - * LISP, ML
 - 逻辑式语言 Logical Language
 - * Prolog
 - 面向对象语言

二、 翻译系统

- 翻译程序 – Translator
 - 将某种语言描述的程序（源程序，Source Code）翻译为等价的另一种语言描述的程序（目标程序，Object Code）的程序。
- 编译程序 – Compiler
 - 将某一种高级语言描述的程序翻译为汇编、机器语言描述的程序
- 解释程序 – Interpreter
- 编译系统 = 编译程序 + 运行系统
- 其他翻译程序
 - 诊断编译程序
 - 优化编译程序
 - 交叉编译程序
 - 可变目标编译程序
 - 并行编译程序
 - 汇编程序、交叉汇编程序、反汇编程序

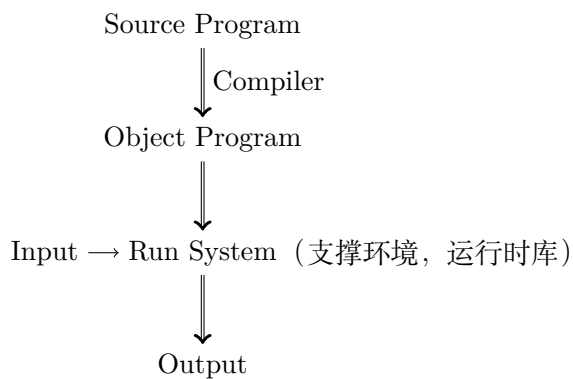


图 1.1: 编译系统 = 编译程序 + 运行系统

三、 翻译系统的功能分析

- 分析
 - 词法，语法，语义
- 翻译
 - 语句的翻译，代码生成
- 例如：标识符左值与右值的绑定
- 变量：存储单元 名字：值
- 函数：目标代码序列 名字：入口地址

四、 翻译系统的总体结构

1.4.1 词法分析

词法分析器 (Lexical Analyzer) 完成词法分析。扫描源程序，并转换为 Token 串，同时查找语法错误，进行标识符登记（管理符号表）。

输入：字符串。
输出：（种类码，属性值）对。

1.4.2 语法分析

语法分析器 (Syntax Analyzer) 完成语法分析。实现『组词成句』，构造分析树，指出语法错误并指导翻译。

输入：Token 序列
输出：语法成分（抽象语法树）

1.4.3 语义分析

语义分析器 (Semantic Analyzer) 分析由语法分析器给出的语法单位的语义。

- 获取标识符的属性：类型、作用域等
- 语义检查：运算合法、取值范围
- 子程序的静态绑定：代码的相对地址
- 变量的静态绑定：数据的相对地址

1.4.4 中间代码生成

升层中间代码，如：前、后缀表达式、三地址表示、LLVM IR。

1.4.5 代码优化

对中间代码的优化处理：提高运行速度、节省储存空间

- 与机器无关的优化
 - 常量合并
 - 公共子表达式提取等
- 与机器有关的优化
 - 循环展开
 - 向量化
 - 访存优化
 - 寄存器排布

1.4.6 目标代码生成

将中间代码转换为目标机器上的指令代码或者汇编代码。



1.4.7 表格管理

管理编译过程中的各种符号表，辅助完成语法、语义检查，完成静态绑定，管理编译过程。

1.4.8 错误处理

进行错误的检查、报告以及纠正。

- 词法错误：拼写、定义；
- 语法错误：语句结构、表达式结构；
- 语义错误：表达式类型不匹配；

1.4.9 前端与后端

将编译过程分为前端与后端。其中：

- 前端：与源语言有关，与目标机器无关的部分
 - 词法分析、语法分析、语义分析、中间代码生成、与机器无关的优化
- 后端：与目标机器有关的部分
 - 与机器有关的代码优化、目标代码生成

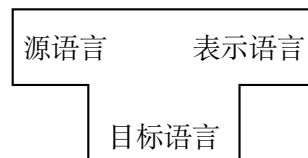
好处：实现新语言和新机器，只需要实现前端、后端中的某一个。

五、 编译程序的生成

通过自动化技术来生成编译程序。

1.5.1 T 形图

用 T 形图表示语言翻译：从源语言翻译为目标语言，而翻译器本身是表示语言。



1.5.2 交叉编译 (Cross Compiling)、移植

例 1.5.1. A 机上有一个 C 语言编译器，是否可以利用此编译器实现 B 机器上的 C 语言编译器？

1. 用 C 语言编制 B 机器的编译程序 P_0 ($C \rightarrow B$)
2. 用 P_1 编译 P_0 ，得到 P_2
3. 用 P_2 编译 P_0 ，得到 P_3

1.5.3 本机编译器的利用

六、 编译技术的应用

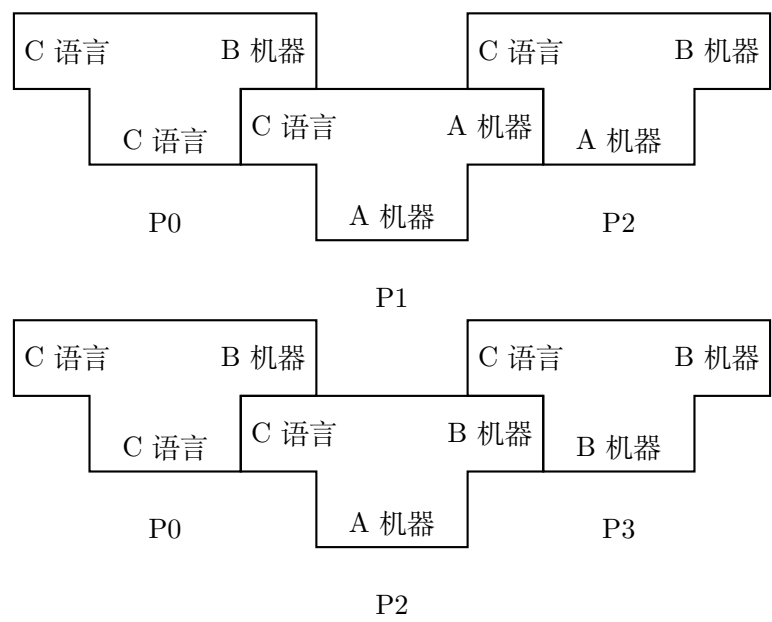


图 1.2: 第 4 页 例 1.5.1 图例