

2021-2022 年度《编译原理上机实习》指导书

一、上机实习目的

理解编译程序的构造原理，掌握编译程序的构造方法与技术。通过实习，使学生既加深对编译原理基础理论的理解，又提高动手能力，特别是提高软件设计能力。

二、上机实习要求

在理解编译原理基本思想的基础上，选择一个自己熟悉的程序设计语言，完成编译程序的设计和实现过程。

编译程序的设计可以采用自顶向下和自底向上两种不同的方法。由于许多高级语言（如 PASCAL，C）中的语法成分都是递归定义的，所以**本实习要求学生采用递归下降分析技术**，这是一种自顶向下的编译方法，其基本思想是对语言的每个（或若干个）语法成分编制一个处理子程序，从处理<程序>这个语法成分的子程序开始，在分析过程中调用一系列过程或函数，对源程序进行语法和语义分析，直到整个源程序处理完毕为止。

本上机实习是为一种架空语言设计一个基于一次扫描的编译程序，完成词法分析、语法分析、语义分析等功能，并生成中间代码(四元式)或某种机器上的目标代码(汇编语言)。

三、上机实习步骤

1. 阅读《上机实习指导书》。
2. 根据设计要求分析并处理文法，写出算法的伪代码，画程序框图
3. 根据程序框图编写编译程序
4. 输入编译程序并上机调试
5. 撰写上机实习报告

四、上机实习内容

1、题目：编译原理上机报告

2、架空语言的文法规则：

<程序> → <变量说明部分>;<语句部分>

<变量说明部分> → <变量说明><标识符列表>

<变量说明> → int

<标识符列表> → <标识符列表>,<标识符>|<标识符>

<标识符> → \$<字母>|<标识符><字母>|<标识符><数字>

<语句部分> → <语句部分><语句>;|<语句>;

<语句> → <赋值语句>|<条件语句>|<循环语句>

<赋值语句> → <标识符>=<表达式>

<条件语句> → if (<条件>) then <嵌套语句>; else <嵌套语句>

<循环语句> → while (<条件>) do <嵌套语句>

<表达式> → <项>|<表达式><加法运算符><项>
 <项> → <因子>|<项><乘法运算符><因子>
 <因子> → <标识符>|<常量>|(<表达式>)
 <常量> → <无符号整数>
 <无符号整数> → <数字序列>
 <数字序列> → <数字序列><数字>|<数字>
 <加法运算符> → +
 <乘法运算符> → *
 <关系运算符> → <|>|!=|>=|<=|==
 <条件> → <表达式><关系运算符><表达式>
 <复合语句> → begin <语句部分> end
 <嵌套语句> → <语句>|<复合语句>
 <字母> → a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z
 <数字> → 0|1|2|3|4|5|6|7|8|9

3、评价基准(合计 115%，超过 100%部分舍去)：

分值分配		要求
词法分析	基本（25%）	从 txt 文件读入源程序，根据词法规则，识别单词，填写相应的表，并在屏幕上输出每次识别出的二元式。如果存在词法错误，则显示错误信息、出错位置
	附加（5%）	出错时试图从错误中恢复。简单的恢复方法是忽略该字符（或单词）或替换成允许的符号后，继续扫描。
语法分析	基本（25%）	对源程序作语法分析，确定是否符合该文法，在屏幕上输出每次推导使用的产生式
	附加（5%）	若存在语法错误，显示错误信息和位置。
语义分析与目标代码生成	语义检查（10%）	在语法分析的基础上，进行语义分析，若存在语义错误（类型匹配错误，重复定义错误），显示错误信息和出错位置。
	代码生成（15%）	往 txt 文件输出源程序的四元式序列；或所学的汇编语言生成目标代码。
上机报告	答辩（10%）	回答问题。
	报告内容（15%）	内容完整，书写字迹工整。
	附加（5%）	第一节上机课完成提交 5%，第二节 4%，第三节 3%，第四节 2%，第五节 1%。

4. 评价方式:

程序: 用源程序生成软件随机产生一个完全符合文法的程序进行4次编译, 包括:

1. 源程序: 词法分析基本 (20) + 语法分析基本 (25) + 代码生成 (15) **(必需)**
2. 修改成带有词法错误的程序: 词法分析基本 (5) + 附加 (5) **(必需)**
3. 修改成带有语法错误的程序: 语法分析附加 (5)
4. 修改成带有语义错误的程序: 语义检查 (10)

答辩: 从题库(上机实习报告中的关于分析方法、算法、思路、总结相关部分的内容)中随机抽取一道问题,

- 若回答无错误且详细 (10);
- 若回答无错误 (5);
- **若回答错误 (0), 并进一步随机抽取词法分析、语法分析、语义分析的基本知识点提问, 回答错误一次对应部分 (-5), 直至回答正确或 0 分为止。**

每个源程序只包含一行语句, **约定单词间无空格** (若源程序生成软件生成的结果中包含空格回车, 可手动或编程删除)。如:

```
int $v, $c; $v=41; $c=15; $c=$v+4*6+($v*3)*$c+940*($v*6); if ($v*89<$v) then $v=$c*479; else begin $c=$c+85; $c=$c+734; end; while ($c*6>$v) do $v=$v+9;
```

五、学生应提交的上机实习材料

完成上机实习后, 每个学生至少应提交如下两种材料:

5.1 上机实习报告 (去教材科领纸质的实习报告, 白色封面的)

要求:

(1) 用钢笔或签字笔填写。不可以用打印稿。

(2) 在实习报告的第三部分(过程部分), 应包含所写编译程序的设计思想、主要的数据结构、主要的函数功能及相关的流程图。

(3) 还应包含测试、运行程序的主要结果。所提供的测试结果, 应能体现编译程序所具有的检错功能。

(4) 实习报告的第四部分(小结部分)除需写明本次实习的收获和体会之外, 还应包括对以下几个问题的讨论: ①在编写调试编译程序的过程中, 遇到了哪些关键问题, 你是如何解决这些问题的。

②在词法分析、语法分析的过程中, 都必须对分析中遇到的错误做出反应, 你的编译程序是如何处理所遇到的错误? 举例说明。

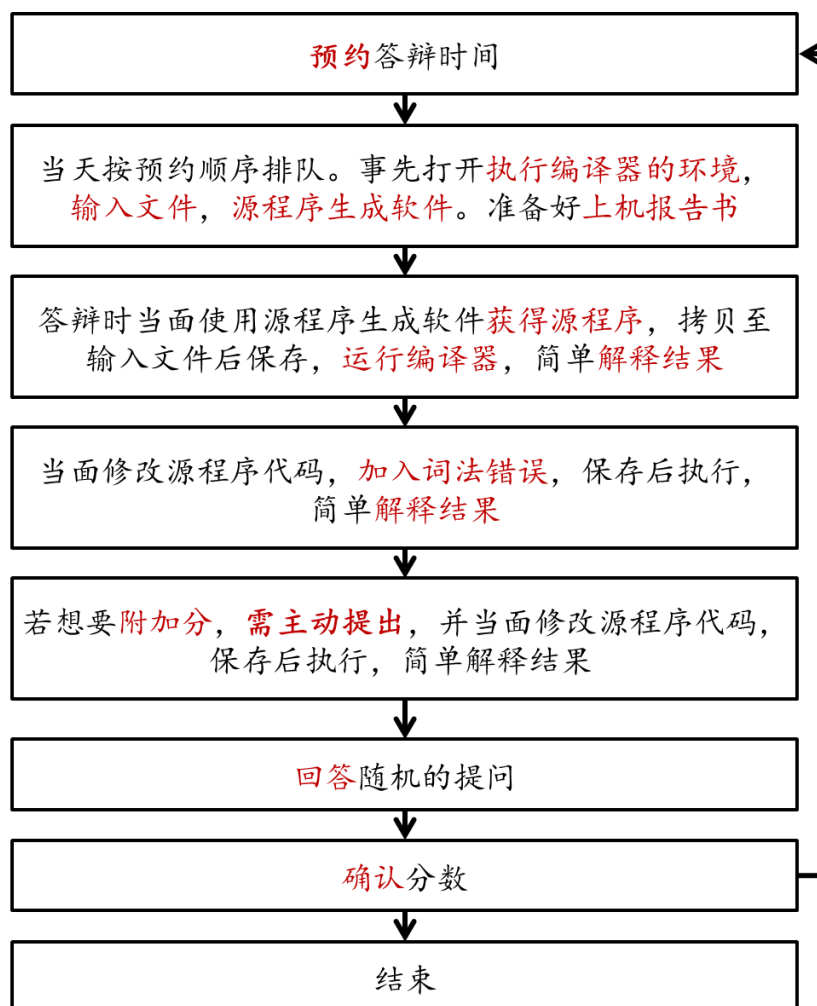
③通常编译程序可以输出被编译的源程序, 假如希望输出优美格式的源程序代码(即, 源程序代码的文本正确缩进, {、}对齐等), 你的编译程序是否实现了此功能? 如何实现的?

5.2 源程序清单及其可执行文件

- (1) 源程序清单，**要求有尽可能详细的注释。**
- (2) 可执行文件(如.exe, .class)应该能在Windows环境中独立执行。
- (3) 输入源程序文件、输出目标代码文件。

5.3 提交方式

上机实习报告：提交流程如下，并当面完成测试，答辩和评分。



注意：

1. **上机课时用于检查结果，编程和撰写报告要利用课余自习时间。**预计每节课90分钟检查30人左右。
2. 留学生无法到校、电脑不带windows系统等**特殊情况主动联系任课老师。**
3. 提交的**报告书不退回。**

**【重要！提交物是否齐全采用程序自动分析判断，
因此文件名需严格按照下述格式命名，文件位置任意】**

程序代码：将源程序文件、可执行文件、运行结果文件压缩打包，提交至qq群作业。

压缩文件名：学号姓名.zip/rar **【注意学号姓名中间无其他符号】**

源程序文件名：学号.c 等

可执行文件名：学号.exe 等

运行结果文件名：in.txt, out.txt **【输入输出文件固定名称】**

目前检查后缀名：.c, .cpp, .cs, .exe, .java., .class, .py, .jar, .txt

- 若源程序和可执行文件用到其他后缀名，在提交作业时标注一下；配置文件无所谓
- 源程序不需要提交完整的项目，只要有可以用文本编辑器打开的、包含了源代码的文件即可

截止日期：最后一节上机课的下课时间。

抄袭和提供抄袭的最终成绩为不及格，并记入档案。

- 全员提交结束后，按年级为单位统一查重

截止日期前**未提交齐程序代码和上机实习报告**的最终成绩最高为**60分**。

材料齐全,但是由于最后一节课时间不够**未完成答辩**的最终成绩最高为**85分(即上机部分零分)**。