

**编译原理实验**

**课程报告**

**TOY语言编译器**

班 级 IS1701

姓 名 李强

学 号 U201714559

指导教师 刘铭

报告日期 2019/12/31

网络空间安全学院

**要 求**

1、报告需本人独立完成，内容真实。引用资料时，需进行标注说明，并列入参考文献中；如发现抄袭，成绩无效；

2、应说明实验的操作系统环境、采用的主要方法、设计的过程、设计的结果（主要源码文件功能、数据结构、函数说明）、遇到的问题、测试运行的情况；

3、按编译原理实验内容，应包含：语言的文法、语言的词法及语法分析、语义分析、中间代码生成、目标代码生成；

4、评分标准：格式规范美观，符合华中科技大学论文格式要求；采用的方法合适、设计合理；4个主要实验环节按任务书要求完成。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **格式**  **规范** | **词法**  **语法** | **语义**  **分析** | **中间**  **代码** | **目标**  **代码** | **总分** |
| **20** | **30** | **30** | **10** | **10** | **100** |
|  |  |  |  |  |  |

**目 录**

[1 选题背景 1](#_Toc28676650)

[1.1 任务 1](#_Toc28676651)

[1.2 目标 1](#_Toc28676652)

[1.3 语言定义 1](#_Toc28676653)

[1.4 主要技术 1](#_Toc28676654)

[2 实验一 词法分析和语法分析 2](#_Toc28676655)

[2.1 单词文法描述（黑体4号加粗, 字母、阿拉伯数字为Time New Roman4号加粗） 2](#_Toc28676656)

[2.2 语法文法描述 3](#_Toc28676657)

[2.3 词法分析器设计 3](#_Toc28676658)

[2.4 语法分析器设计 3](#_Toc28676659)

[2.5 词法及语法分析器实现结果 3](#_Toc28676660)

[2.6 小结 3](#_Toc28676661)

[3 实验二 语义分析 4](#_Toc28676662)

[3.1 语义表示方法描述 4](#_Toc28676663)

[3.2 符号表结构 4](#_Toc28676664)

[3.3 错误类型定义 4](#_Toc28676665)

[3.4 语义分析实现技术 4](#_Toc28676666)

[3.5 语义分析结果 4](#_Toc28676667)

[3.6小结 4](#_Toc28676668)

[4 实验三 中间代码生成 5](#_Toc28676669)

[4.1 中间代码格式定义 5](#_Toc28676670)

[4.2 中间代码生成规则定义 5](#_Toc28676671)

[4.3 中间代码生成过程 5](#_Toc28676672)

[4.4 代码优化（可选） 5](#_Toc28676673)

[4.5 中间代码生成结果 5](#_Toc28676674)

[4.6 小结 5](#_Toc28676675)

[5 实验四 目标代码生成 6](#_Toc28676676)

[5.1 指令集选择 6](#_Toc28676677)

[5.2 寄存器分配算法 6](#_Toc28676678)

[5.3 目标代码生成算法 6](#_Toc28676679)

[5.4 目标代码生成结果 6](#_Toc28676680)

[5.5 目标代码运行结果 6](#_Toc28676681)

[5.6 小结 6](#_Toc28676682)

[6 总结 7](#_Toc28676683)

[参考文献 8](#_Toc28676684)

# 1 选题背景

1.1 任务

通过简单自定义语言编译器的完整实现，掌握编译原理理论知识，提高灵活运用理论知识以解决实际问题的能力；提高系统软件编写能力。

1.2 目标

课程目标是构造一个高级语言的子集的编译器，目标代码是汇编语言。通过设计自定义语言的文法，进行词法分析、语法分析和语义分析最终生成中间代码IR来实现自定义语言的前端。通过设计目标机器架构寄存器调度算法，指令调度方法等实现自定义语言的后端。最后通过链接器将汇编代码转换为机器码。

1.3 语言定义

本次实验中实现的语言为TOY。

TOY语言采用了C语言中的int，float和char数据类型，函数定义和函数原型声明，while循环和if流程控制等。采用Python语言中以换行表示一条语句结束的方法，结合C语言中以大括号开始的代码域，语言中不采用分号结尾。

G[program]:

program → ExtDefList//开始

ExtDefList → ε //变量定义或函数定义或函数声明

| Specifier DecList ENTERS ExtDefList // 外部变量

| Specifier DecList ExtDefList //外部变量

| Specifier FunDec CompSt ENTERS ExtDefList //函数声明

| Specifier FunDec ENTERS ExtDefList

Specifier → int | float | char//类型

DecList → Dec | Dec , DecList //解决一次性定义多个情况

Dec → VarDec | VarDec = Exp //初始化问题

VarDec → ID//变量名字对应到ID

FunDec → ID ( VarList ) //形参比较特殊并没有分配实际的空间因此需要特殊处理

VarList → ε//无参

| ParamDec

| ParamDec , VarList //参数变量定义，同样是递归定义

ParamDec → Specifier VarDec//参数只是临时或内部变量

CompSt → { ENTERS DefList StmList }//代码段

DefList → ε | Specifier DecList ENTERS DefList

StmList → ε | Stmt StmList

Stmt → Exp ENTERS

| CompSt ENTERS

| return Exp ENTERS

| if ( Exp ) ENTERS CompSt ENTERS

| if ( Exp ) ENTERS CompSt else CompSt ENTERS

| while ( Exp ) ENTERS CompSt ENTERS

| BREAK ENTERS

Exp → Exp =Exp | Exp += Exp | Exp -= Exp | Exp \*= Exp | Exp /= Exp //复合赋值运算

| ++ Exp | -- Exp | Exp -- | Exp ++ //自增自减运算

| ! Exp | Exp && Exp | Exp || Exp//逻辑运算

| Exp < Exp | Exp <= Exp | Exp == Exp | Exp != Exp | Exp > Exp | Exp >= Exp //比较运算

| Exp + Exp | Exp – Exp | Exp \* Exp | Exp / Exp | Exp % Exp //算术运算

| ( Exp ) //括号表达式

| - Exp

| ID ( Args )

| ID

| INT | FLOAT |CHAR //终结符

Args→ ε

| Exp , Args

| Exp

ENTERS→ ENTER

| ENTER ENTERS

1.4 主要技术

编译器所在环境为Linux。通过Flex和Bison编写词法分析器和语法分析器并生成抽象语法树，通过对抽象语法树进行递归遍历进行语义分析。因为TAC序列没有一个完整的标准，并且对于与或非等逻辑表达式没有注明，对于比较运算并没有转换，这不利于翻译到汇编代码，因此采用同样是静态单赋值三地址码形式的LLVM的IR[1]。在实现TOY编译器后端时，可以借助LLVM的框架将TOY前端整合到LLVM中，或是自写特定机器架构的后端或是直接使用LLVM已经实现的部分后端，构建完整的编译器。

# 2 实验一 词法分析和语法分析

2.1 单词文法描述

然后2-4章，介绍你的解决方案和细节。章节数可以适当调整

2.2 语法文法描述

2.3 词法分析器设计

2.4 语法分析器设计

2.5 词法及语法分析器实现结果

2.6 小结

3 实验二 语义分析

3.1 语义表示方法描述

3.2 符号表结构

3.3 错误类型定义

3.4 语义分析实现技术

3.5 语义分析结果

3.6小结

4 实验三 中间代码生成

4.1 中间代码格式定义

4.2 中间代码生成规则定义

4.3 中间代码生成过程

4.4 代码优化（可选）

4.5 中间代码生成结果

4.6 小结

5 实验四 目标代码生成

5.1 指令集选择

5.2 寄存器分配算法

5.3 目标代码生成算法

5.4 目标代码生成结果

5.5 目标代码运行结果

5.6 小结

6 总结

--------说明：章与章之间应插入分页符，但每章结束页，空白部分不要超过2/3，如最后一页大面积空白，可以将内容适当调整；报告双面打印；模板章节标题可根据自己实验内容适当调整----------

参考文献(黑体小2号加粗居中)(最后列出你的参考文献)

[1] 王静康,张凤宝,夏淑倩等.论化工本科专业国际认证与国内认证的“实质性”.高等工程教育研究,2014,5:1-4 http://llvm.org/docs/LangRef.html

[2] Stone J A, Howard L P. A simple technique for observing periodic nonlinearities in Michelson interferometers. Precision Engineering,1998,22(4):220-232

[3] 朱印红,袁衍明.Dreamweaver完美网页设计——技术入门篇.(第一版).北京:中国电力出版社,2006:19～20

[4] Lewis S L. Physics and chemistry of the solar system.北京:北京大学出版社,2014.1～2

[5]□陈剑.上博简《民之父母》“而得既塞於四海矣”句解释[EB/OL］.简帛研究网站，http://www.bamboosilk.org/Wssf/2003/chenjian03.htm．2003-01-18

( 宋体小4号)

……

……

……