

**实践课程报告**

**题目： C--语言编译器**

**课程名称： 编译原理实践**

**专业班级： CS1409**

**学 号： U201414797**

**姓 名： 张丹朱**

**指导教师： 邵志远**

**报告日期： 2017年1月14日**

**计算机科学与技术学院**

**目录**

[**1选题背景** 1](#_Toc472160273)

[1.1 任务 1](#_Toc472160274)

[1.2 目标 1](#_Toc472160275)

[1.3 源语言定义 1](#_Toc472160276)

[**2实验一 词法分析和语法分析** 2](#_Toc472160277)

[2.1 单词文法描述 2](#_Toc472160278)

[2.2 语言文法描述 2](#_Toc472160279)

[2.3 词法分析器的设计 2](#_Toc472160280)

[2.4 语法分析器设计 2](#_Toc472160281)

[2.5 语法分析器实现结果展示 2](#_Toc472160282)

[**3实验二 语义分析** 3](#_Toc472160283)

[3.1 语义表示方法描述 3](#_Toc472160284)

[3.2 符号表结构定义 3](#_Toc472160285)

[3.3 错误类型码定义 3](#_Toc472160286)

[3.4 语义分析实现技术 3](#_Toc472160287)

[3.5 语义分析结果展示 3](#_Toc472160288)

[**4结束语** 4](#_Toc472160289)

[4.1 实践课程小结 4](#_Toc472160290)

[4.2 自己的亲身体会 4](#_Toc472160291)

[**参考文献** 5](#_Toc472160292)

# 1选题背景

1.1 任务

主要是通过对简单编译器的完整实现，加深课程中关键算法的理解，提高学生对系统软件编写的能力。

1.2 目标

本次课程实践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。按照任务书，实现的方案可以有很多种选择。

1.3 源语言定义

可以根据自己对编程语言的喜好选择实现。建议大家选用C语言的简单集合C--语言或教材中的Decaf语言。

本次实验选择C--语言作为源语言。

# 2实验一 词法分析和语法分析

2.1 单词文法描述

使用正则表达式来进行单词文法的描述，在C--中基本的单词及其正则表达式如下表1.1所示。

（其中，附加了对八进制和十六进制数以及指数形式的浮点数的识别。

表1.1 单词正则表达式表

|  |  |
| --- | --- |
| 单词 | 正则表达式 |
| INT |  |
| FLOAT |  |
| ID |  |
| SEMI |  |
| COMMA |  |
| ASSIGNOP |  |
| RELOP |  |
| PLUS |  |
| MINUS |  |
| STAR |  |
| DIV |  |
| AND |  |
| OR |  |
| DOT |  |
| NOT |  |
| TYPE |  |
| LP |  |
| RP |  |
| LB |  |
| RB |  |
| LC |  |
| RC |  |
| STRUCT |  |
| RETURN |  |
| IF |  |
| ELSE |  |
| WHILE |  |

另外，对于单行和多行注释也进行了识别，将注释内容全部忽略。

2.2 语言文法描述

2.3 词法分析器的设计

2.4 语法分析器设计

2.5 语法分析器实现结果展示

# 3实验二 语义分析

3.1 语义表示方法描述

在实验一中生成的语法树的基础上为各结点增加相应的属性，主要是增加类型。各语法单元类型的确定

然后利用符号表和语法树结点本身的属性，对相应的短语中的单元进行类型的匹配。

3.2 符号表结构定义

3.3 错误类型码定义

本次实验，除了完成实验书上给出的所有错误类型外，另外增加错误类型18和19。最后所有错误类型的定义如下表3.1所示

3.4 语义分析实现技术

3.5 语义分析结果展示

# 4结束语

4.1 实践课程小结

本次实验主要完成了如下几点工作：

（1）确定了所有单词的正则表达式，实现基本单词识别，附加实现八进制和十六进制数的识别、指数形式的浮点数的识别和单行、多行注释的识别。

（2）确定了语法树的结构，生成了语法树。

（3）完成了对实验一的调试与测试工作。

（4）确定了符号表的结构，实现了符号表的相关操作函数，实现了对整棵语法树的语义分析。

（5）完成了对语义分析的调试与测试工作。

4.2 自己的亲身体会

总的来说，本次实验任务设计的比较合理，若是能够把整个编译器完全实现，那么对于理解编译原理相关理论知识会有极大的帮助。唯一不足的是实验的时间安排并不合理，安排的时间太少了，跟不上进度。

对于词法和语法分析，使用了flex和bison工具之后节省了大量的时间，有利于后面工作的展开，但同时失去了一次对于有限状态机和分析表的构造与使用的练习机会，好在利用工具时可以看到其生成的源程序，能够帮助我理解和体会词法分析与语法分析的过程。

在语法分析构造语法树的时候构造了一棵孩子兄弟表示法的二叉树，使用的时候发现很不方便，构造成多叉树会好一些。

在语义分析的时候动态维护符号表，一出作用域就将该作用域的相关符号信息从符号表中删去，这样在语义分析的时候降低了空间的消耗，方便了对作用域所处层数的判断，使每层的作用域清楚得区分开来，在语义分析过程中有很大帮助。但是在后续生成中间代码的过程中这样的符号表使得没法将生成中间代码的程序单独一个文件，只能在语法分析的同时进行词法分析，有诸多不便，下次在设计符号表的时候需要好好考虑如何使得符号表能够被高效地利用。

另外，像错误类型、语法单元等数据都可以用枚举，方便使用，做实验的时候考虑不周，比较是否是某种类型时较为复杂。。

总之，虽然没有完整地完成实验，但是掌握了flex和bison的使用，熟悉了语法树和符号表的构造及使用，练习了树的生成与遍历、链表的创建和查找，对于自顶向下的语义分析也有了较深的体会，收获颇丰。

**参考文献**

[1] 吕映芝等. 编译原理(第二版). 北京：清华大学出版社，2005

[2] 胡伦俊等. 编译原理(第二版). 北京：电子工业出版社，2005

[3] 王元珍等. 80X86汇编语言程序设计. 武汉：华中科技大学出版社,2005

[4] 王雷等. 编译原理课程设计. 北京：机械工业出版社，2005

[5] 曹计昌等. C语言程序设计. 北京：科学出版社，2008