

Syntax FrontEnd Design

Chengyuan ZHANG

March 2021

任 务 书

设计内容

在计算机科学中，抽象语法树 (*abstract syntax tree* 或者缩写为 *AST*)，是将源代码的语法结构的用树的形式表示，树上的每个结点都表示源程序代码中的一种语法成分。之所以说是“抽象”，是因为在抽象语法树中，忽略了源程序中语法成分的一些细节，突出了其主要语法特征。抽象语法树 (*Abstract Syntax Tree*, *AST*) 作为程序的一种中间表示形式，在程序分析等诸多领域有广泛的应用。利用抽象语法树可以方便地实现多种源程序处理工具，比如源程序浏览器、智能编辑器、语言翻译器等。

在《高级语言源程序格式处理工具》这个题目中，首先需要采用形式化的方式，使用巴克斯 (*BNF*) 范式定义高级语言的词法规则 (字符组成单词的规则)、语法规则 (单词组成语句、程序等的规则)。再利用形式语言自动机的原理，对源程序的文件进行词法分析，识别出所有单词；使用编译技术中的递归下降语法分析法，分析源程序的语法结构，并生成抽象语法树，最后可由抽象语法树生成格式化的源程序。[2] [1]

设计要求

1. 语言定义

选定 C 语言的一个子集，要求包含：

- (1) 基本数据类型的变量、常量，以及数组。不包含指针、结构，枚举等。
- (2) 双目算术运算符 (+ - * / %)，关系运算符、逻辑与、逻辑或、赋值运算符。不包含逗号运算符、位运算符、各种单目运算符等等。
- (3) 函数定义、声明与调用。
- (4) 表达式语句、复合语句、if 语句的 2 种形式、while 语句、for 语句，return 语句、break 语句、continue 语句、外部变量说明语句、局部变量说明语句。
- (5) 编译预处理 (宏定义，文件包含)
- (6) 注释 (块注释与行注释)

2. 单词识别

设计 DFA 的状态转换图 (参见实验指导)，实验时给出 DFA，并解释如何在状态迁移中完成单词识别 (每个单词都有一个种类编号和单词的字符串这 2 个特征值)，最终生成单词识别 (词法分析) 子程序。

注：含后缀常量，以类型不同作为划分标准种类编码值，例如 123 类型为 int，123L 类型为 long，单词识别时，种类编码应该不同；但 0x123 和 123 类型都是 int，种类编码应该相同。

3. 语法结构分析

- (1) 外部变量的声明；
- (2) 函数声明与定义；
- (3) 局部变量的声明；
- (4) 语句及表达式；
- (5) 生成 (1)-(4)（包含编译预处理和注释）的抽象语法树并显示。

4. 按缩进编排生成源程序文件。

5. 评测说明

要求具有如下功能：

(1) 识别语言的全部单词（50%）：

要求测试用例包含所有种类的单词，测试用例中没有出现的单词种类视作没有完成该类单词的识别。由于每类单词有一个种类编码（参见实验指导书用枚举常量定义），可以将识别出来的单词按种类编码进行排序显示，这样既能方便自己的调试，也能方便检查。注意相同种类编码的多种形式，都应该包含在测试用例中，例如类型为 int 的常量，有三种形式 0123、123、0x123。

报错功能，指出不符合单词定义的符号位置。

(2) 语法结构分析与生成抽象语法树（40%）：

要求测试用例包含函数声明，定义、表达式（各种运算符均在某个表达式中出现）、所有的语句，以及 if 语句的嵌套，循环语句的嵌套。测试用例中没有出现的语句和嵌套结构，视作没有完成该种语法结构的分析。

报错功能，指出不符合语法规则的错误位置。测试文件中不必包含错误语句等，检查时由老师随机修改测试文件，设置错误，检查报错功能是否实现。

显示抽象语法树，要求能由抽象语法树说明源程序的语法结构，这也是检查时验证语法结构分析正确性的依据。

(3) 缩进编排重新生成源程序文件（10%）：

对（2）的测试用例生成的抽象语法树进行先根遍历，按缩进编排的方式写到.c 文件中，查看文件验证是否满足任务要求。

参考文献

- [1] 严蔚敏等. 数据结构 (*C* 语言版). 北京: 清华大学出版社, 1999.
- [2] 王生原, 董渊, 张素琴, 吕映芝等. 编译原理 (第 3 版). 北京: 清华大学出版社, 2010.

目录

任务书	1
第一章 引言	5
1.1 背景与研究意义	5
1.2 国内外研究现状	5
1.3 课程设计的主要研究工作	5
第二章 系统需求分析与总体设计	6
2.1 系统需求分析	6
2.2 系统总体设计	6
第三章 系统详细设计	7
3.1 有关数据结构的定义	7
3.2 主要算法设计	7
第四章 系统实现与测试	8
4.1 系统实现	8
4.2 系统测试	8
第五章 结论	9
5.1 讨论与总结	9
5.2 未来展望	9
第六章 心得体会	10

第一章 引言

1.1 背景与研究意义

1.2 国内外研究现状

1.3 课程设计的主要研究工作

第二章 系统需求分析与总体设计

2.1 系统需求分析

2.2 系统总体设计

```
1  #include<stdio.h>
2  int main(){
3  }
```

第三章 系统详细设计

3.1 有关数据结构的定义

3.2 主要算法设计

第四章 系统实现与测试

4.1 系统实现

4.2 系统测试

第五章 结论

5.1 讨论与总结

5.2 未来展望

第六章 心得体会