

《编译原理实验》说明

一、实验文档与成绩说明

1、文件提交

- (1) 提交网址: <http://172.22.112.254>, 用户名: byyl, 无密码
- (2) 截止时间: 2017 年 12 月 30 日 12: 00
- (3) 提交文件要求: 每位同学对每一个实验均要提交电子版和纸质报告, 以及实验源程序。首先提交电子版文件至上述网址, 电子版文件包括源程序(必须有)、实验报告(必须有), 测试用例、输出文件等。请确保提交的源程序能编译运行, 否则不要上交。纸质报告在电子版及源程序检查完成后由班级收集后再**统一交**到助教处。纸质报告以问题分析、介绍所采用数据结构及算法和主要函数功能、实验中遇到的问题及解决方法、实验测试(含屏幕截图)、实验心得等为主, 不得大量引用源程序(引用源程序总行数不得超过 100 行), 可参见实验报告模板。

2、成绩评定

实验报告占最终期评成绩 40%, 三个实验算平均得分。程序运行检查、讲解、现场问题处理等占期评成绩 60%。

不得互相拷贝, 一经发现, 雷同的实验均会判为不及格。

3、实验时间

15 周周二、周三(12-12, 12-13) 答疑与程序检查

17 周周二、周三(12-26, 12-27) 程序检查

地点: 计算中心三楼机房(暂定 B302)

二、实验题目及内容

实验一: 正则表达式与有限自动机

实验目的: 通过本次实验, 加深对正则表达式、NFA、DFA 及其识别的语言的理解

实验题目: 以下题目任选一

1、编程实现 DFA 识别字符串的过程。

- 实验内容:**
- (1) DFA 的输入与 DFA 的存储: 确定 DFA 的数据结构以及存储格式
 - (2) DFA 的正确性检查: DFA 的五元组是否正确
 - (3) 输入任意一个整数 N, DFA 的能列表显示其识别的所有长度小于等于 N 的字符串;
 - (4) DFA 的规则字符串判定: 输入(或用字符集随机生成)一个字符串, 模拟 DFA 识别字符串的过程判定该字符串是否是规则字符串(属于 DFA 的语言集)。

2、编程实现正则表达式转换为 NFA, 并将 NFA 确定化为 DFA

实验内容: 输入任意一个正则表达式, 输出其对应的 NFA。正则表达式中运算符只有连接 \cdot 、选择 $|$ 和闭包 $*$ 以及括号。确定 NFA 的存储格式、要求输出 NFA 的五元组各信息, 并对 NFA(五元组)进行确定化, 输出等价的 DFA。

***3、**结合 1-3, 设计一个字符串的识别器。任意给定一个正则表达式, 输入任意的字符串, 判定该字符串是否属于给定的正则语言集

***4、**编程实现网页源代码分析, 能从网页中提取各种标签、网址等信息(信息种类可自己定)。要求首先实现网络爬虫, 从网络中抓取一定数量的网页

实验二 自顶向下语法分析

实验目的：通过本次实验，加深对 LL(1)分析法中 FIRST 与 FOLLOW 集合的计算以及 LL(1)文法的判定算法与分析过程的理解，学习程序设计语言的语法分析器的手工编程方法。

实验题目：以下题目任选一

1、递归下降分析程序

实验内容：(1) 将递归下降子程序分析方法运用于算术表达式文法；

* (2) 生成如图所示的抽象语法树(Abstract Syntax Tree)，设计并实现一个简单的计算器。

2.LL(1)分析程序

2、LL(1)分析法的实现

实验内容：对任意给定一个 CFG，判断文法是否存在左递归和左公因子，如果有，则先进行改写，然后计算文法的 FIRST 集合、FOLLOW 集合、构造 LL(1)分析表，并判定文法是否是 LL(1)文法。

扩展要求：LL(1)分析过程。

实验三 由底向上语法分析

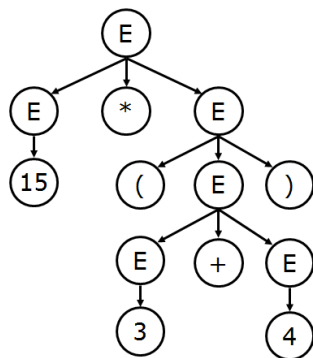
实验目的：通过本次实验，加深对移进规约分析法中算符优先法及 LR(0)的理解，学习程序设计语言的语法分析器的手工编程方法。

1、算符优先分析法

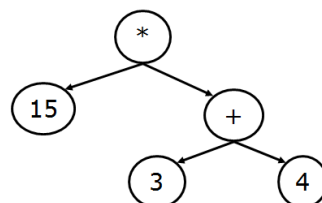
实验题目：算符优先分析法的编程实现，

实验内容：(1) 对于任意给定的文法，判断其是否是算符优先文法，如果是则构造出算符优先关系表。

* (2) 通过语法分析，生成如图所示的抽象语法树，在此基础上设计并实现一个简单的计算器。



Parse tree



Abstract syntax tree

2、LR 分析

实验题目：构造识别活前缀的 DFA

实验内容：输入一个 CFG，编程构造其识别活前缀的 DFA，并判断是否有冲突。要求输出 DFA 的转换函数或状态转换图，以及 LR(0)项目集规范簇。