编译原理实验报告：词法分析器

15331416 赵寒旭 数字媒体技术

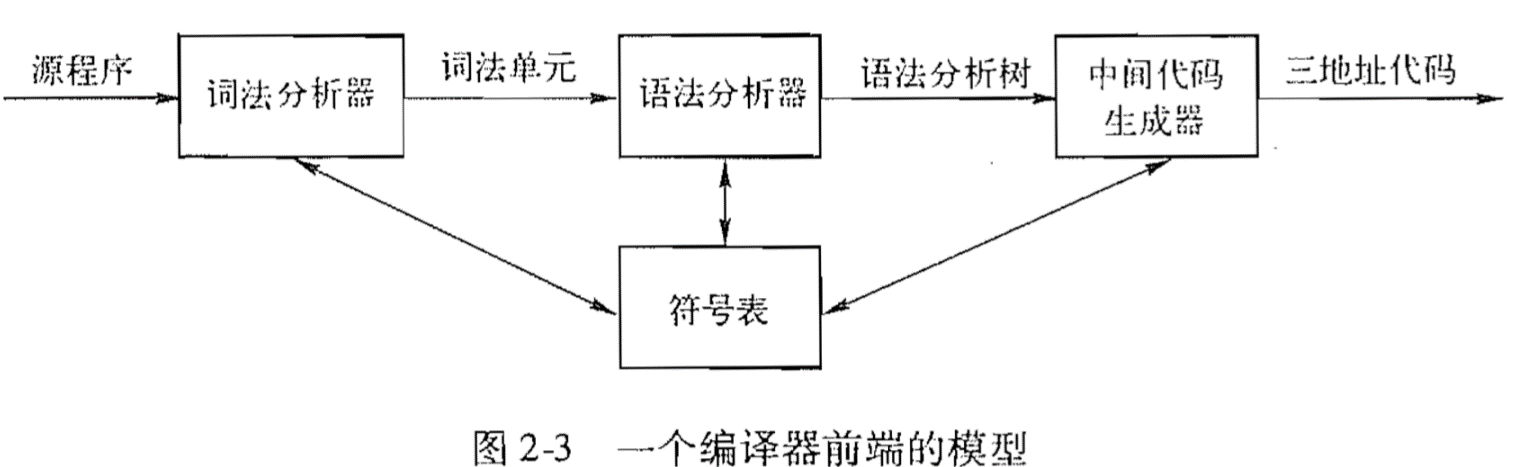
1. 实验目的

用lex实现一个词法分析器

2. 实验思路

2.1 词法分析器的作用

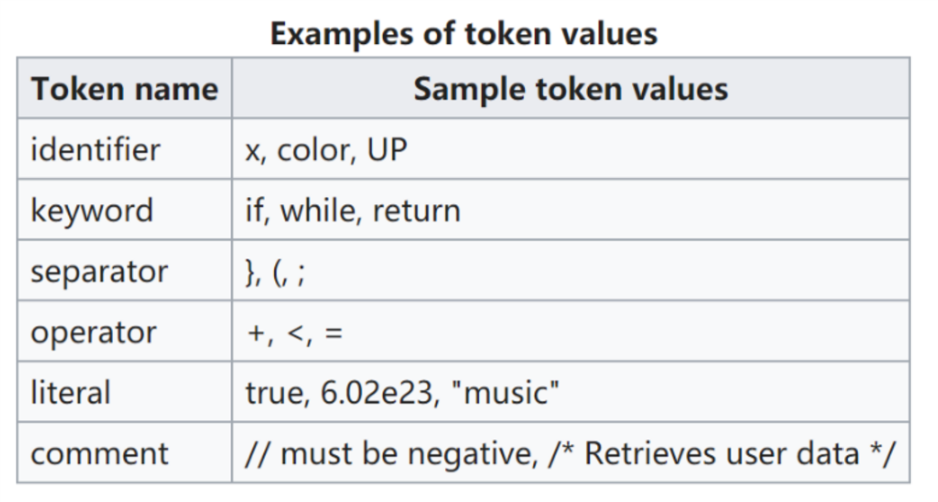
词法分析是编译的第一阶段，主要任务是读入源程序的输入字符，将它们组成词素，生成并输出一个词法单元序列，每个词法单元对应一个词素。本次实验中词法分析程序使用lex自动生成。



2.2 词法单元类型

引自<https://en.wikipedia.org/wiki/Lexical_analysis#Token>

本文词法单元类型及内容参考下表：



identifier（标识符）：由程序员选择的名称

keyword（关键字）：已由编程语言确定的名称

separator（分隔符）：标点符号和成对的分隔符号

operator（操作符）：操作参数产生结果

literal（文字量）：数字和文本

comment（注释）：包括行注释和块注释

2.3 词法分析器生成工具lex

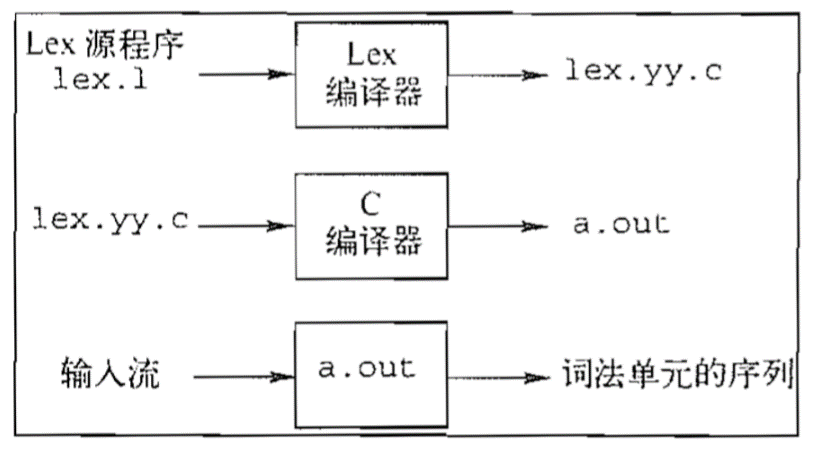
lex支持使用正则表达式来描述各个词法单元的模式，由此给出一个词法分析器的规约。

（1）用lex创建一个词法分析器的过程

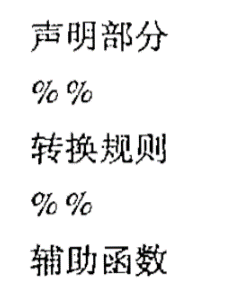
1) 用Lex语言写一个输入文件lex.l，描述将要生成的词法分析器。

2）Lex编译器将lex.l转换成c语言程序lex.yy.c

3）lex.yy.c被c编译器便以为一个.exe文件，得到一个读取输入字符流并生成词法单元流的可运行的词法分析器。



（2）Lex程序的结构



1）声明部分

首先可有起始于"%{"符号，终止于"%}"符号包括include语句、声明语句在内的C语句。

包括变量和明示常量，可以声明词法单元的名字或正则定义。

2）转换规则

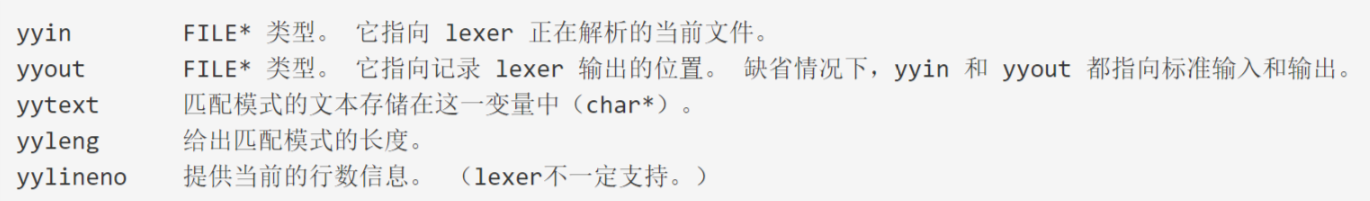
形式：模式 { 动作 }

每个模式是一个正则表达式，可以使用声明部分中定义的变量。

3）辅助函数

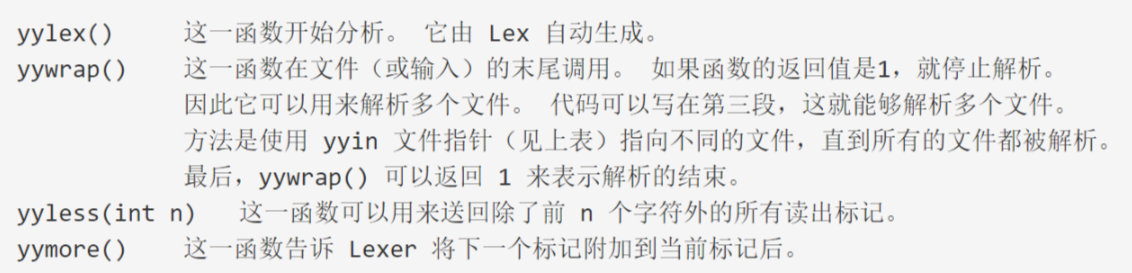
包含各个动作需要使用的所有辅助函数。

2.4 Lex变量



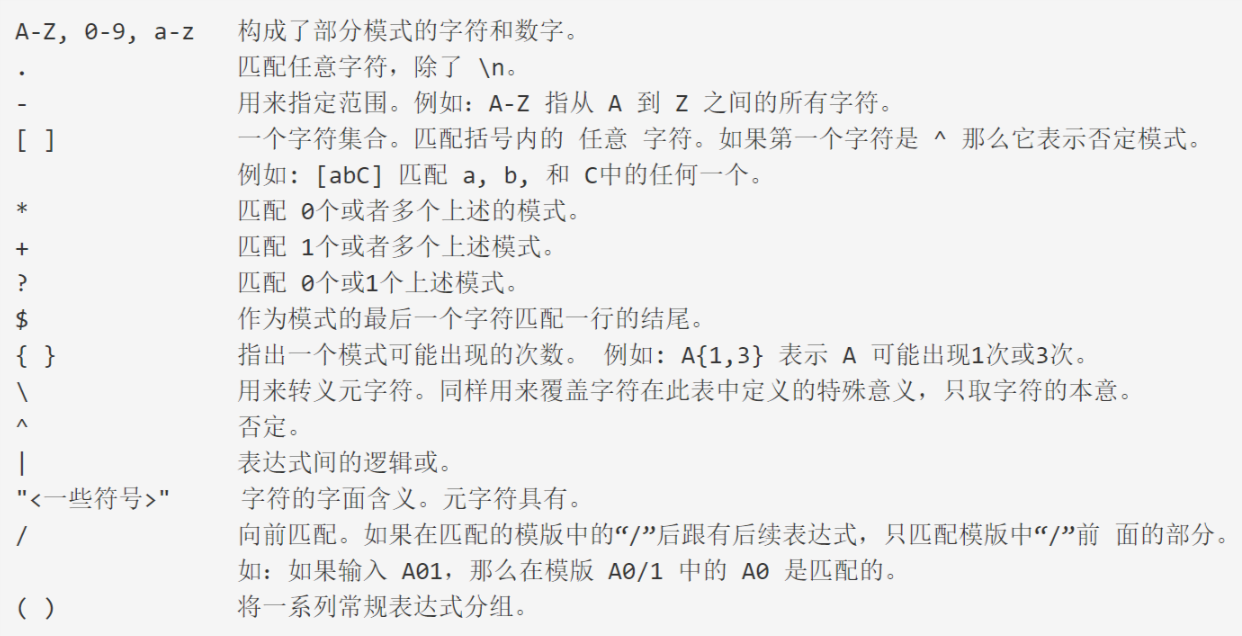
实验中使用yyin指向待解析代码文件，用yytext获得模式匹配文本

2.5 Lex函数



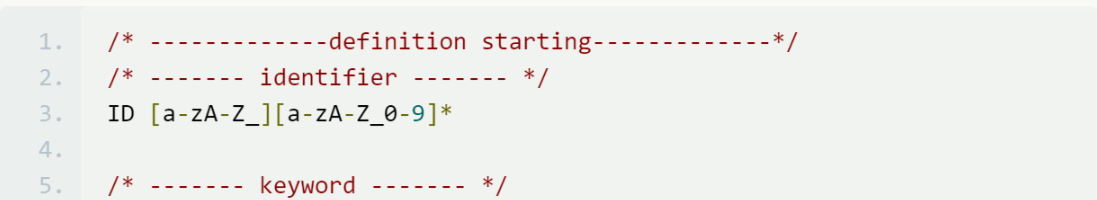
实验中使用yylex()函数分析代码文本，使yywrap()返回值为1解析完毕停止解析。

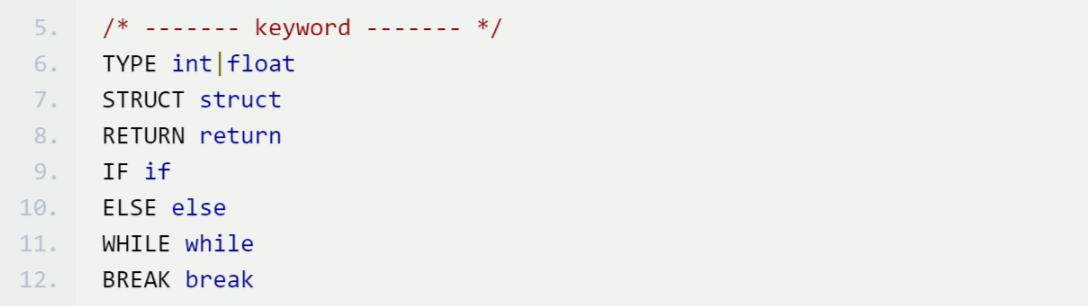
2.5 正则表达式相关字符及其含义列表

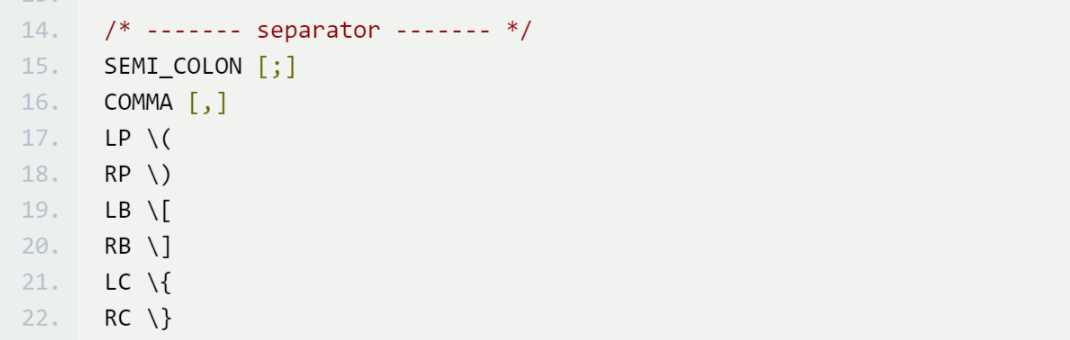


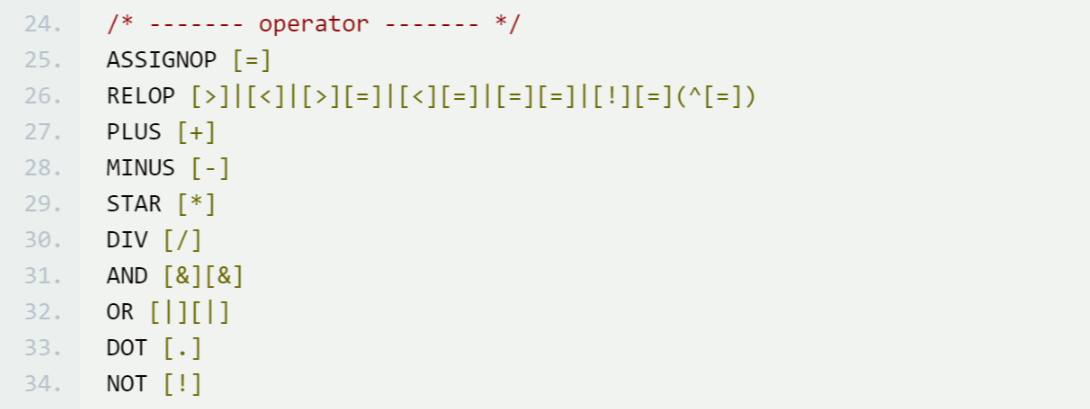
3. 实验步骤

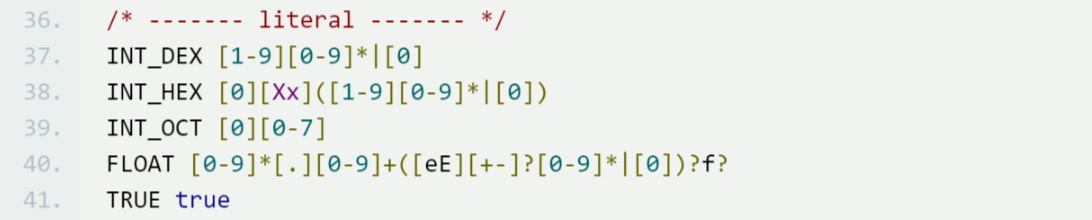
3.1 定义需要匹配的具体词法单元对象

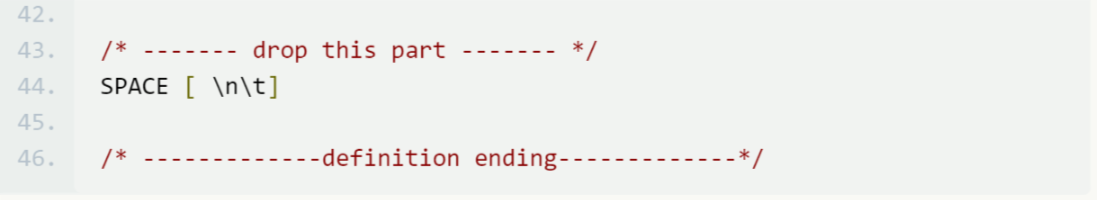












3.2 匹配后操作的实现

3.3

4. 实验结果

4.1 运行步骤

4.2 运行结果

5. 遇到的问题

Lex中的冲突解决

选择在Lex程序中先被列出的模式，id应该放在后面！