

《编译原理》实验要求

实验时间：第9-13周，周三下午5-8节，其中第13周只有1节，共17课时

实验地点：线上

一、词法分析器的设计(必做题)

编程实现一个 C 语言子集的词法分析器，要求至少包含以下功能：

- (1) 实现以字母或_等符号打头的标识符的识别；
- (2) 实现无符号整数或无符号数的识别；
- (3) 实现保留字的识别；
- (4) 实现算术运算符的识别；
- (5) 实现关系运算符的识别；
- (6) 实现分隔符的识别；
- (7) 实现非法字符的报错；
- (8) 构造一个保留字表 `reslist`，并事先将保留字存入其中；
- (9) 构造一个用户标识符表 `idlist`，包含标识符名、标识符类型、存储长度等字段。

词法分析器运行过程中将源程序中出现的标识符（包括用户变量名、用户函数名等）填入标识符名字段，标识符类型字段、存储长度字段等为空。

注意：源程序中用户标识符出现多次的在用户标识符表中只对应 1 项。

- (10) 构造一个无符号整数的常数表 `uintlist`，包含数值、类型、存储长度等字段。

词法分析器运行过程中将源程序中出现的无符号整数填入数值字段，类型、存储长度等字段为空。注意：源程序中出现多次的整数在无符号整数表中只对应 1 项。

- (11) 词法分析器从 `input` 文件中读入一小段 C 语言源程序的语句，以二元式的形式按顺序输出其所有单词。输出结果在显示器上显示，同时存入 `output` 文件。C 语言源程序中所有的用户标识符都要添加到用户标识符表 `idlist` 中，所有的无符号整数都要添加到常数表 `uintlist` 中。
- (12) 程序的输出文件至少包括：`output`, `idlist`, `uintlist`，其中 `output` 文件中的单词必须以二元式的形式存储。

实验评分时查看如下内容：

- (1) 源代码
- (2) 输入信息：`input` 文件、保留字表 `reslist`
- (3) 输出信息：`output`, `idlist`, `uintlist`

二、语法分析器的设计(选做题)

编程构造有优先级的、无左递归的算术表达式文法 $G[E]$ 的 LL(1) 分析表，要求至少包含如下子程序：

- (1) 计算文法 $G[E]$ 的 FIRST 集；
- (2) 计算文法 $G[E]$ 的 FOLLOW 集；
- (3) 构造文法 $G[E]$ 的 LL(1) 分析表

$G[E]: E \rightarrow TA$

$A \rightarrow +TA \mid \varepsilon$

$T \rightarrow FB$

$B \rightarrow *FB \mid \varepsilon$

$F \rightarrow (E) \mid i$