

# labLexer 实验报告

---

学号：19373249 姓名：邱学思

## 简要介绍

这次词法分析小实验的整体算法构造我参考了课本《编译原理及编译程序构造》p55 图 3.7 的代码。在主函数中循环调用 getsym() 函数直至读取到文件末尾或遇到不符合 token 规则的字符串/字符。

对于单次 getsym() 函数的调用：

1. 跳过空白符：首先读取文件当前的字符，若该字符为空格、换行符或者制表符，则循环读取直至遇到第一个非空白符或者已经读取到文件末尾。
2. 判断是否为标识符或关键字：若当前字符属于 Nondigit，根据标识符的文法循环读取字符直至遇到既不属于 Nondigit 也不属于 Digit 的字符或者读取到文件末尾，此时将文件的指针回退 1，接着判断拼接后的 token 是否属于关键字并输出（最长匹配原则：在符合文法的前提下先尽可能多地拼接，再判断是否完全匹配上关键字）。
3. 判断是否为无符号整数：若当前字符属于 Digit，根据无符号整数的文法循环读取字符直至遇到不属于 Digit 的字符或者读取到文件末尾，此时将文件的指针回退 1，输出拼接完成后的 token。
4. 其他类型判断：若均不满足，则输出 `Err` 并退出程序。

## 思考题回答

本次实验无思考题