

Lab2

尚晴宇

How you handle comments:

增加COMMENT状态，当遇到/*时进入该状态，对于除/*，*/外的其他符号全部忽略。考虑到嵌套注释的可能，C声明中维护了一个注释层数变量，初始为0。当遇到*/时加1，遇到/*时减1，减1后判断层数是否为0，如果为0就回到INITIAL状态，代表注释结束。

How you handle strings:

增加STR状态，当遇到“时进入该状态，在变量stringPos中保存字符串开始的位置，在str中保留字符串内容，通过strcat函数追加来维护，再次遇到”时返回，返回时设置开始位置及字符串内容。针对字符串中可能出现的各种转义字符，匹配时要注意在正则表达式中加\代表后面的字符为字符本身。按照课本附录提示对字符串进行转换，追加不同内容。当遇到其中\f__f\时，增加状态STRIGN来忽略其中的格式化字符。

error handling:

错误处理是判断程序中是否有不属于正则表达式的字符串，在INITIAL下加对.的处理，当所有都无法匹配时调用EM_error，输出错误位置（第几行，第几字符）和错误信息——非法字符。

EOF handling:

当整个文件正常结束遇到EOF时不需处理，当在字符串中或注释中未遇到终止符（/*/)而文件结束时，调用EM_error报错，表示字符串注释不完整。

Other interesting features of your lexer:

功能上没有什么特殊的地方，在string完成过程中遇到小bug，最终通过新建char数组存放内容并在末尾填0然后用于追加解决了问题。熟悉errorMsg等其他文件，应用了EM_error并在C声明中进行封装，注意错误提示时位置的维护。