

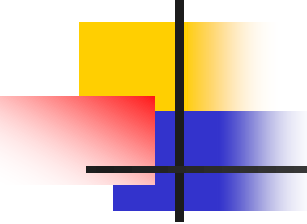


词法分析作业简介



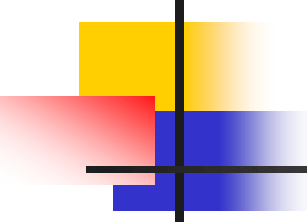
目标和要求：

- 根据具体文法编写词法分析程序，考核学生对词法分析方法的掌握情况，培养学生编写符合词法规则的词法分析程序的能力。
- 学生在理论课所学词法分析方法的基础上，设计实现词法分析程序，并按规定的格式输出单词信息
- 作业提交至教学平台，用5个测试程序进行测试，并进行相似性检查。根据输出结果与预期结果一致部分所占的比例给分。



请根据给定的文法设计并实现词法分析程序，从源程序中识别出单词，记录其单词类别和单词值，输入输出及处理要求如下：

- 数据结构和与语法分析程序的接口请自行定义；类别码需按下表格式统一定义；
- 为了方便进行自动评测，约定：
 - 输入的被编译源文件统一命名为`testfile.txt`；
 - 输出的结果文件统一命名为`output.txt`；
 - 结果文件中每行按如下方式组织：
单词类别码 单词的字符/字符串形式(中间仅用一个空格间隔)



单词的类别码请统一按如下形式定义：

单词名称	类别码	单词名称	类别码	单词名称	类别码	单词名称	类别码
标识符	IDENFR	if	IFTK	-	MINU	=	ASSIGN
整型常量	INTCON	else	ELSETK	*	MULT	;	SEMICN
字符常量	CHARCON	do	DOTK	/	DIV	,	COMMA
字符串	STRCON	while	WHILETK	<	LSS	(LPARENT
const	CONSTTK	for	FORTK	<=	LEQ)	RPARENT
int	INTTK	scanf	SCANFTK	>	GRE	[LBRACK
char	CHARTK	printf	PRINTF TK	>=	GEQ]	RBRACK
void	VOIDTK	return	RETURNTK	==	EQL	{	LBACE
main	MAINTK	+	PLUS	!=	NEQ	}	RBRACE

例如有如下程序段：

```
const int const1 = 1, const2 = -100;
const char const3 = '_';
int change1;
char change3;
int gets1(int var1,int var2){
    change1 = var1 + var2;
    return (change1);
}
```

则输出的结果文件具有的内容为：

```
CONSTTK const
INTTK int
IDENFR const1
ASSIGN =
INTCON 1
COMMA ,
IDENFR const2
ASSIGN =
MINU -
INTCON 100
SEMICN ;
CONSTTK const
CHARTK char
IDENFR const3
ASSIGN =
CHARCON _
SEMICN ;
INTTK int
IDENFR change1
SEMICN ;
CHARTK char
IDENFR change3
SEMICN ;
```

```
INTTK int
IDENFR gets1
LPARENT (
INTTK int
IDENFR var1
COMMA ,
INTTK int
IDENFR var2
RPARENT )
LBACE {
IDENFR change1
ASSIGN =
IDENFR var1
PLUS +
IDENFR var2
SEMICN ;
RETURNTK return
LPARENT (
IDEN change1
RPARENT )
SEMICN ;
RBACE }
```