

西北工业大学

《编译原理》词法分析实验

学 院：	_____ 软件学院 _____
学 号：	_____ 2018303081 _____
姓 名：	_____ 马泽红 _____
专 业：	_____ 软件工程 _____

西北工业大学

2021 年 4 月

目录

1. 测试方案.....	3
1.1 手动测试.....	3
1.2 自动测试.....	3
1.3 枚举变量说明.....	3
2. 手动测试.....	4
2.1 标识符（ID）测试.....	4
2.2 整数（Integer）测试.....	4
2.3 字符串（String）测试.....	5
2.4 布尔型（Boolean）测试.....	5
2.5 转义字符测试.....	6
2.6 专有符号测试.....	6
2.7 保留字（Reserved Word）测试.....	7
3. 自动测试.....	7
3.1 标识符（ID）测试.....	7
3.2 表达式（Exp）测试.....	8
3.3 复杂结构测试.....	8
3.4 保留字测试.....	9
3.5 专有符号测试.....	9

1. 测试方案

本代码参考 Tiny 编译器编写而成，主要使用了双层 Switch-case 完成有限状态机代码的翻译与编写。

同时，状态机和代码中均包含了对字符串中**转义字符**的识别。

1.1 手动测试

通过控制台手动输入设计好的测试代码，观察代码输出是否与预期一致。运行代码后，输入数字 1 即可进入手动测试模式，然后便可输入自测代码。

```
请选择测试方式
1:控制台手动测试。
2:文件自动读写测试:

1
控制台测试
mazehong;
mazehong                                ID
;                                         SPECIFICSYMBOL
```

1.2 自动测试

程序自动加载提供的五个大类的测试文件，然后自动生成词法分析结果并存放在 output 文件夹中。运行代码后输入数字 2 即可自动测试。

```
请选择测试方式
1:控制台手动测试。
2:文件自动读写测试:

2
自动测试
E:\电子书\编译原理\2021编译原理词法分析程序撰写的实验作业4.10\compiler_experiment1\MIDL_compiler\src\test\测试文件\专有符号\test_3_formula.txt自动测试结束
E:\电子书\编译原理\2021编译原理词法分析程序撰写的实验作业4.10\compiler_experiment1\MIDL_compiler\src\test\测试文件\关键字\test_1_kw.txt自动测试结束
E:\电子书\编译原理\2021编译原理词法分析程序撰写的实验作业4.10\compiler_experiment1\MIDL_compiler\src\test\测试文件\关键字\test_2_kw.txt自动测试结束
E:\电子书\编译原理\2021编译原理词法分析程序撰写的实验作业4.10\compiler_experiment1\MIDL_compiler\src\test\测试文件\复杂结构\test_8_complex.txt自动测试结束
E:\电子书\编译原理\2021编译原理词法分析程序撰写的实验作业4.10\compiler_experiment1\MIDL_compiler\src\test\测试文件\复杂结构\test_9_complex_with_escape.txt自动测试结束
```

1.3 枚举变量说明

```
public enum TokenType{

    ID,                //标识符

    NUM,               //整数
```

```
    STRING,          //字符串
    BOOLEAN,         //布尔型
    RESERVEDWORD,    //保留关键字
    SPECIFICSYMBOL,  //专用符号
    ERROR            //错误
}
```

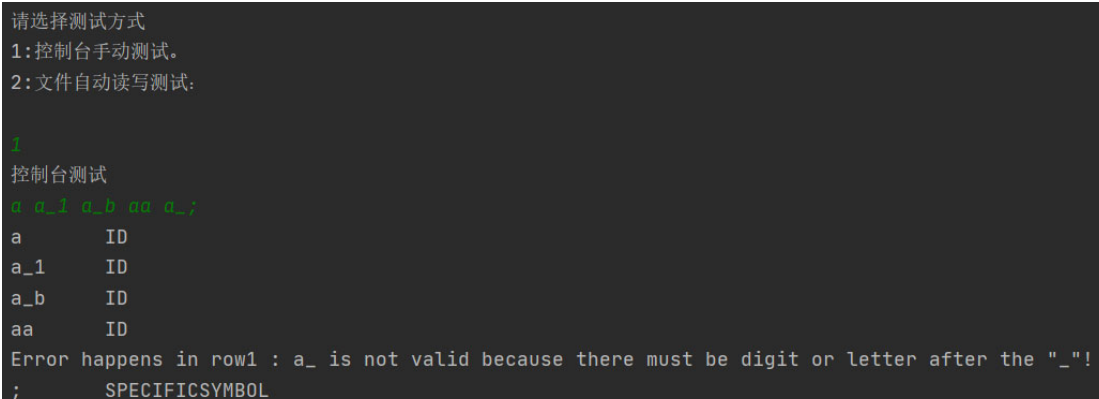
2. 手动测试

所有测试用例均为自己根据路径覆盖科学设计，完全覆盖了 DFA 中的所有路径，验证了代码实现的完整性和正确性。

2.1 标识符（ID）测试

通过控制台手动输入以下测试用例，覆盖 DFA 中标识符的状态转移路径，观察输出是否与预期一致。

```
Start-> 1 -> Done : a
Start-> 1 -> 2 -> 1 -> Done : a_1 | a_b
Start-> 1 -> 1 -> Done : aa
Start-> 1 -> 2 -> Done : a_!
```



```
请选择测试方式
1:控制台手动测试。
2:文件自动读写测试:

1
控制台测试
a a_1 a_b aa a_!
a      ID
a_1    ID
a_b    ID
aa     ID
Error happens in row1 : a_ is not valid because there must be digit or letter after the "_"!
;      SPECIFICSYMBOL
```

2.2 整数（Integer）测试

```
Start-> 3 -> Done : 0; | 01
```

Start-> 3 -> 5 -> Done : 01 | 0L

Start-> 4 -> 5 -> Done : 1L;

Start-> 4 -> 4 -> Done : 12;

控制台测试

0; 01 0L 0L 1L 12

0 NUM

; SPECIFICSYMBOL

Error happens in row1 : 0 because zero can't be followed by digit

0L NUM

0L NUM

1L NUM

12 NUM

2.3 字符串 (String) 测试

Start-> 6 -> 6 -> 8 -> Done : “a”

Start-> 6 -> 7 -> 6 -> 8 -> Done : “\t” | “\n” | “\f”

Start-> 6 -> 7 -> Done : “a\m”

状态机和代码中均加入了对转义字符的识别。

请选择测试方式

1: 控制台手动测试。

2: 文件自动读写测试:

1

控制台测试

"a" "\t" "\n" "\f"

"a" STRING

"\t" STRING

"\n" STRING

"\f" STRING

"a\m"

Error happens in row1 : "a\m escape character is not valid!

2.4 布尔型 (Boolean) 测试

Start -> 9 -> Done: True | False

```
请选择测试方式
1:控制台手动测试。
2:文件自动读写测试:

1
控制台测试
False True
False          BOOLEAN
True           BOOLEAN
```

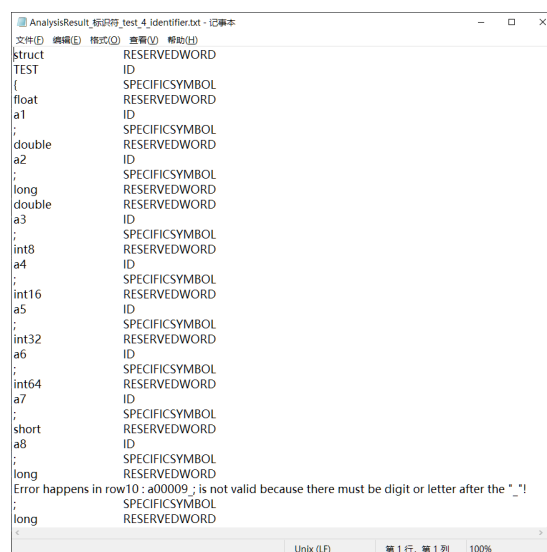
2.7 保留字 (Reserved Word) 测试

```
请选择测试方式
1:控制台手动测试。
2:文件自动读写测试:

|
控制台测试
"struct", "float", "boolean", "short", "long", "double",
        "int8", "int16", "int32", "int64", "uint8", "uint16", "uint32",
        "uint64", "char", "unsigned" "struct"      STRING
,
SPECIFICSYMBOL
"float"      STRING
,
SPECIFICSYMBOL
"boolean"    STRING
,
SPECIFICSYMBOL
"short"      STRING
,
SPECIFICSYMBOL
"long"       STRING
,
SPECIFICSYMBOL
"double"     STRING
,
SPECIFICSYMBOL
"int8"       STRING
,
SPECIFICSYMBOL
"int16"      STRING
,
SPECIFICSYMBOL
"int32"      STRING
```

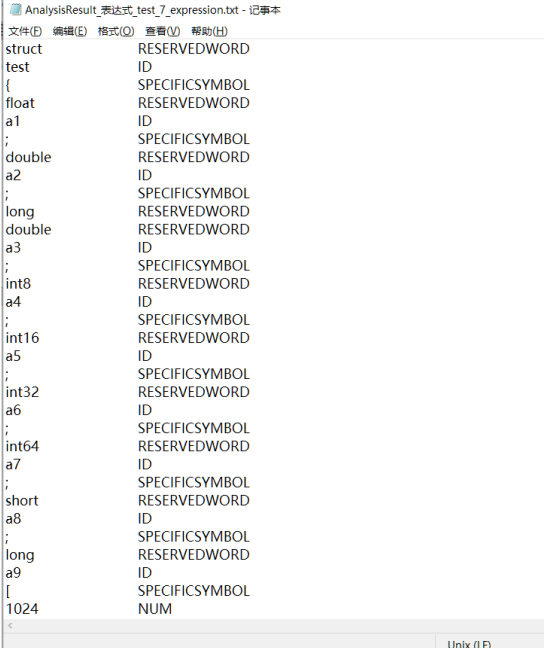
3. 自动测试

3.1 标识符 (ID) 测试



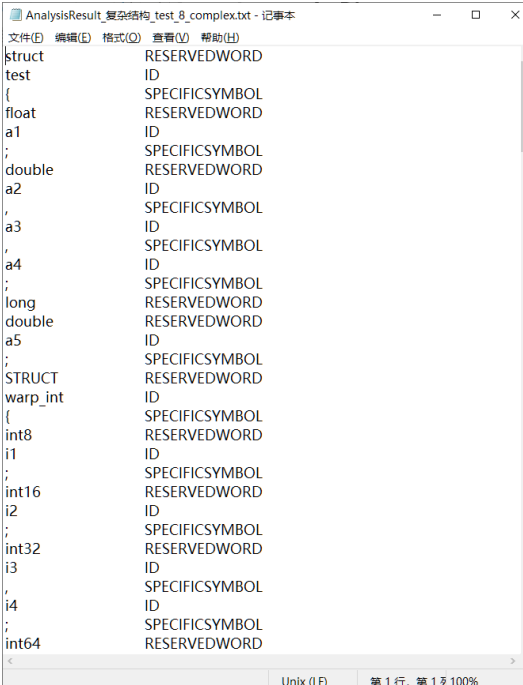
```
AnalysisResult_标识符_test_4_identifier.txt - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
struct      RESERVEDWORD
TEST        ID
{           SPECIFICSYMBOL
float       RESERVEDWORD
a1          ID
;           SPECIFICSYMBOL
double      RESERVEDWORD
a2          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
double      RESERVEDWORD
a3          ID
;           SPECIFICSYMBOL
int8         RESERVEDWORD
a4          ID
;           SPECIFICSYMBOL
int16        RESERVEDWORD
a5          ID
;           SPECIFICSYMBOL
int32        RESERVEDWORD
a6          ID
;           SPECIFICSYMBOL
int64        RESERVEDWORD
a7          ID
;           SPECIFICSYMBOL
short        RESERVEDWORD
a8          ID
;           SPECIFICSYMBOL
long         RESERVEDWORD
Error happens in row10 : a00009; is not valid because there must be digit or letter after the \".'!
;           SPECIFICSYMBOL
long         RESERVEDWORD
```

3.2 表达式（Exp）测试



```
AnalysisResult_表达式_test_7_expression.txt - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
struct      RESERVEDWORD
test        ID
{           SPECIFICSYMBOL
float       RESERVEDWORD
a1          ID
;           SPECIFICSYMBOL
double      RESERVEDWORD
a2          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
double      RESERVEDWORD
a3          ID
;           SPECIFICSYMBOL
int8        RESERVEDWORD
a4          ID
;           SPECIFICSYMBOL
int16       RESERVEDWORD
a5          ID
;           SPECIFICSYMBOL
int32       RESERVEDWORD
a6          ID
;           SPECIFICSYMBOL
int64       RESERVEDWORD
a7          ID
;           SPECIFICSYMBOL
short       RESERVEDWORD
a8          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
a9          ID
[           SPECIFICSYMBOL
1024       NUM
```

3.3 复杂结构测试



```
AnalysisResult_复杂结构_test_8_complex.txt - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
struct      RESERVEDWORD
test        ID
{           SPECIFICSYMBOL
float       RESERVEDWORD
a1          ID
;           SPECIFICSYMBOL
double      RESERVEDWORD
a2          ID
;           SPECIFICSYMBOL
a3          ID
;           SPECIFICSYMBOL
a4          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
double      RESERVEDWORD
a5          ID
;           SPECIFICSYMBOL
STRUCT      RESERVEDWORD
warp_int    ID
{           SPECIFICSYMBOL
int8        RESERVEDWORD
i1          ID
;           SPECIFICSYMBOL
int16       RESERVEDWORD
i2          ID
;           SPECIFICSYMBOL
int32       RESERVEDWORD
i3          ID
;           SPECIFICSYMBOL
i4          ID
;           SPECIFICSYMBOL
int64       RESERVEDWORD
```


3.4 保留字测试

AnalysisResult_关键字_test_1_kw.txt - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
struct      RESERVEDWORD
test        ID
{           SPECIFICSYMBOL
float       RESERVEDWORD
a1          ID
;           SPECIFICSYMBOL
double      RESERVEDWORD
a2          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
double      RESERVEDWORD
a3          ID
;           SPECIFICSYMBOL
int8        RESERVEDWORD
a4          ID
;           SPECIFICSYMBOL
int16       RESERVEDWORD
a5          ID
;           SPECIFICSYMBOL
int32       RESERVEDWORD
a6          ID
;           SPECIFICSYMBOL
int64       RESERVEDWORD
a7          ID
;           SPECIFICSYMBOL
short       RESERVEDWORD
a8          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
a9          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
```

< Unix (LF)

3.5 专有符号测试

AnalysisResult_专有符号_test_3_formula.txt - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
struct      RESERVEDWORD
test        ID
{           SPECIFICSYMBOL
float       RESERVEDWORD
a1          ID
;           SPECIFICSYMBOL
double      RESERVEDWORD
a2          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
double      RESERVEDWORD
a3          ID
;           SPECIFICSYMBOL
int8        RESERVEDWORD
a4          ID
;           SPECIFICSYMBOL
int16       RESERVEDWORD
a5          ID
;           SPECIFICSYMBOL
int32       RESERVEDWORD
a6          ID
;           SPECIFICSYMBOL
int64       RESERVEDWORD
a7          ID
;           SPECIFICSYMBOL
short       RESERVEDWORD
a8          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
a9          ID
;           SPECIFICSYMBOL
long        RESERVEDWORD
```

< Unix (LF)