

Project 1

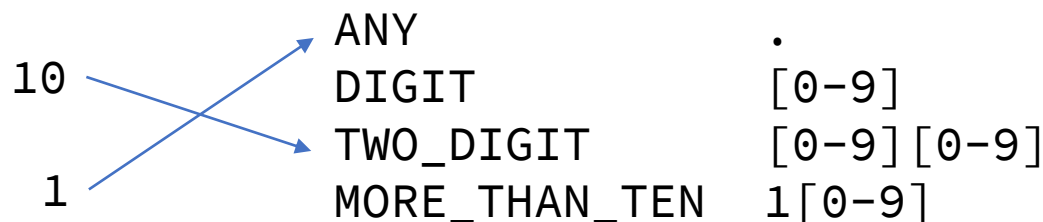
Q&A

- Q: 如何输出到文件?

- A: `printf()`

- `./proj1 < hello.mjava > hello.out`

- Q: 注释里有单引号', 或者字符串里有/*...*/, 如何处理?
- A: 如果同时满足.l文件中的多个正则表达式规则, Lex会匹配其中**长度最长** (被匹配的字符数最多) 的规则, 如果同时满足两个以上的长度相同的规则, 则最先定义的规则会被匹配。



```
'a /*comment*/ within a string'  
/* a 'string' within a comment */
```

无需特殊处理, 不影响Lex识别

- Q: 如何识别并输出错误?

- A: `// lexer.l`

```
ID          [a-zA-Z][a-zA-Z0-9]*  
WRONG_ID     [0-9][a-zA-Z0-9]*
```

```
{WRONG_ID} {  
    error = "Error: Malformed identifier '" + yytext + "'";  
    return(ERROR);  
}
```

```
// driver.c
```

```
if (lexRtn == ERROR) {  
    printf("%s, at line %d, column %d\n", error, yyline, yycolumn);  
}
```

- Q: 输出结果中的Index in String table可以是字符串的Index, 而不是字符的Index吗?

- A: 不可以

```
/* Example 1: A hello world program */
program xyz;
class Test {
    method void main() {
        System.println('Hello World !!!');
    }
}
```

Line	Column	Token	Index in String table
2	8	PROGRAMnum	
2	12	IDnum	0
2	13	SEMIum	
3	6	CLASSnum	
3	11	IDnum	4
3	13	LBRACEnum	
4	11	METHODnum	
4	16	VOIDnum	
4	21	IDnum	9
4	22	LPARENnum	
4	23	RPARENnum	
4	25	LBRACEnum	
5	15	IDnum	14
5	16	DOTnum	
5	23	IDnum	21
5	24	LPARENnum	
5	41	SCONSTnum	29
5	42	RPARENnum	
5	43	SEMIum	
6	6	RBRACEnum	
7	2	RBRACEnum	
8	1	EOFnum	

String Table : xyz Test main System println Hello World !!!

- Q: ICONST需要放进String Table里吗?

- A: 不用

- Q: `yywrap()`是用来做什么的?
- A: 当扫描完一个文件后, Lex会调用`yywrap`函数查询是否还有新的文件需要识别, 返回0表示有新文件, 返回1表示没有新文件。

```
int yywrap() {  
    return 1;  
}
```


或

```
%option noyywrap
```

其他注意事项

- 函数需要先声明后使用

```
%{  
#include "token.h"  
  
/* TODO: declare variables here */  
  
%}  
  
/* TODO: define regular expressions */  
ELSE      (else)  
  
/* TODO: define translation rules: pattern and action */  
%%  
{ELSE}    { return (ELSEEnum); }  
  
%%  
  
/* TODO: define the auxiliary functions you may use */  
  
int yywrap() {  
    return 1;  
}
```



其他注意事项

- 提交：把所有文件放到一个文件夹里，然后压缩成.tar.gz文件
- 评分：

条目	分数
1. 程序编写 (25%)	
(1) 结构, 注释, 清晰度 (15%)	
(2) 结果输出 (10%)	
2. 功能实现 (75%)	
(1) 忽略大小写 (Case insensitivity) (5%)	
(2) 识别正确的标识符 (IDs) (5%)	
(3) 识别正确的字符串常量 (string constants) (5%)	
(4) 识别正确的关键字 (keywords) (5%)	
(5) 识别转义序列 (escape sequences) (5%)	
(6) 识别其他 token (other tokens) (5%)	
(7) 识别正确的注释 (comments) (5%)	
(8) Same ID/string same index (7.5%)	
(9) Same string same index (7.5%)	
(10) Detect not ending strings (7.5%)	
(11) Detect not ending comments (7.5%)	
(12) Detect wrong IDs (5%)	
(13) Continue after errors (5%)	
总计 (100 分)	