

自动糖果售货机编程作业

2020-11

主要内容

- **Lex**简介
- **Lex**源文件的格式
- **Lex**安装及编程环境设置
- 自动糖果售货机及其**DFA**
- 作业要求

Lex简介

- 词法分析（lexical analysis）是计算机科学中将字符序列转换为单词（Token）序列的过程。进行词法分析的程序或者函数叫作**词法分析器**（Lexical analyzer），也叫**扫描器**（Scanner）。
- 词法分析器一般以函数的形式存在，供语法分析器调用。
- Lex和Yacc是美国Bell实验室用C语言研制的词法分析程序和语法分析程序的自动生成工具。它们的强大功能使创建高级语言（如，C语言）的编译器十分轻松。
- 该工具与C语言强耦合。

Lex的工作流程

LEX源程序
lex.l → LEX编译器 → Lex.yy.c

lex.yy.c → C编译器 → Lex.yy.o

字符流源程序 → a.out → 记号序列

- Flex的输入文件称为Lex源文件，它内含正规表达式和对相应模式处理的C语言代码。Lex源文件的扩展名习惯上用.l表示。
- Flex通过对源文件的扫描自动生成相应的词法分析函数 `int yylex()`，并将之输出到名规定为 `lex.yy.c` 的文件中。该文件即为Lex的输出文件或输出的词法分析器。
- 应用中可将 `yylex()` 加入自己的工程文件中使用。

Lex源文件的格式

- 一个Lex源程序由三部分组成，以%%分隔：
- 定义部分（定义全局变量）
%%
- 规则部分（定义正规表达式以及相应的动作）
%%
- 用户附加C语言（C语言主程序main以及附加的处理程序）

下面以统计单词出现的次数的源程序count.l为例：

定义部分

包括：(1). 由：分隔符“% {”和“% }”扩起的C代码；
(2). 正则表达式的名字定义；

%{ /* yytext是Lex 赋予并由正则表达式匹配的串的内部
名字， * /

#include <stdlib.h>

#include <stdio.h>

int count = 0;

%}

INTEGER [1-9][0-9]*

ID [a-zA-Z][a-zA-Z_0-9]*

SPACE [/n/t]

% %

... ..

C程序代码

名字定义

规则部分

- `%%`
`{INTEGER} { num_num++;`
`printf("(num=%d)/n",atoi(yytext));/*打印数字值*/`
`/*数字数加一*/`
`}`

`{ID} { num_id++;`
`printf("(id=%s)/n",yytext);`
`}`

`{SPACE} |`
`.`
`{. /*可匹配任何字符 */`
`/*什么也不做,滤掉白字符和其它字符*/`
`}`

`%%`

用户附加程序部分

- %%

```
int main()
{
    yylex();
    printf("num=%d, id=%d/n", num_num, num_id);
    return 0;
}
```

```
int yywrap() //此函数必须由用户提供
```

```
{
    return 1;
}
```

Lex正则表达式的元字符列表如下：

元字符	匹配内容
.	除了换行符之外的任意字符
\n	换行符
*	0 次或者多次匹配
+	1 次或者多次匹配
?	0 次或者 1 次匹配
^	行首
\$	行尾
a b	a 或者 b
(ab)+	ab 的一次或者多次匹配
“a+b”	a+b（字面意思）
[]	一类字符

Lex编程环境安装与设置

开发Lex和Yacc 程序最需要的程序就是lex和yacc了，如果你是Unix 或者Linux 系统，则系统自带了这两个工具。在Windows 系统环境下的Lex是flex，而Yacc 则是bison，需要安装以下环境：

1. Lex(flex.exe)和 Yacc(bison.exe)环境
2. C/C++编译器

这些软件包是开放的，下载网址为：

<http://unxutils.sourceforge.net/>

<https://sourceforge.net/projects/flex/>（64位）

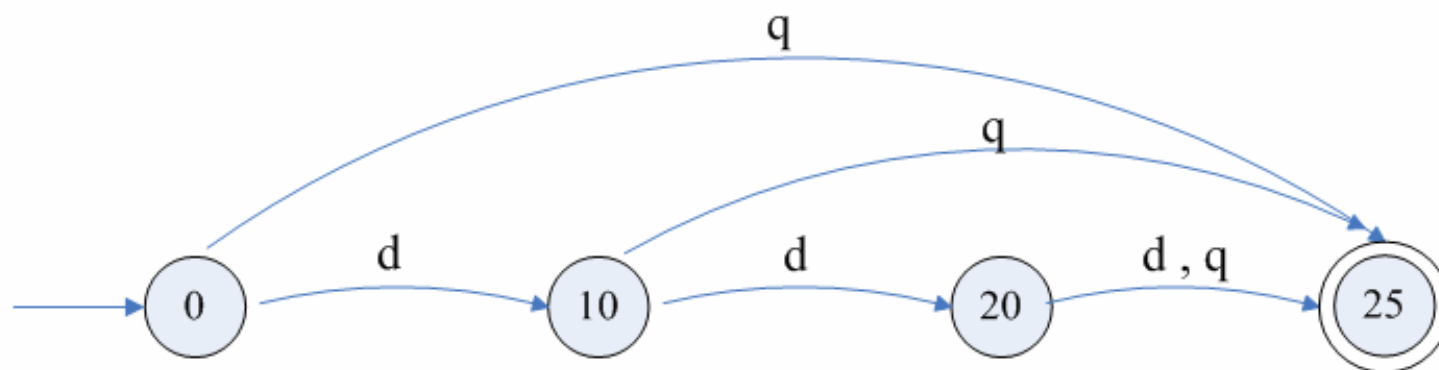
Lex编程环境安装与设置

下载flex.rar（本课程同学可去BB平台获取该软件包），解压缩后，运行Setup即可安装flex-2.5.4a-1。安装完毕后，要设置Windows系统的PATH环境变量,增加flex.exe文件所在的目录（如：安装在c:\program files\GnuWin32）：
右键点击“计算机”、“属性”、“高级系统设置”、“环境变量”，在下面系统变量里面找到PATH，修改，在后面加上

c:\program files\GnuWin32\bin

现在就可以使用flex了，找一个目录，把你的test.1文件（flex文件以*.1为后缀）放进去，然后输入flex test.1，就会出现一个lex.yy.c了。

自动糖果售货机及其DFA（简化版）



d=dime（10美分），q=quarter(25美分)

Fig. 自动糖果售货机的DFA

使用图上作业法可求得其对应的正则表达式为：

$dd(d+q)^+(dq+q)$

用Lex中的正则表达式符号表示为：

$dd[dq] | dq | q$

自动糖果售货机的**lex.l**文件及运行

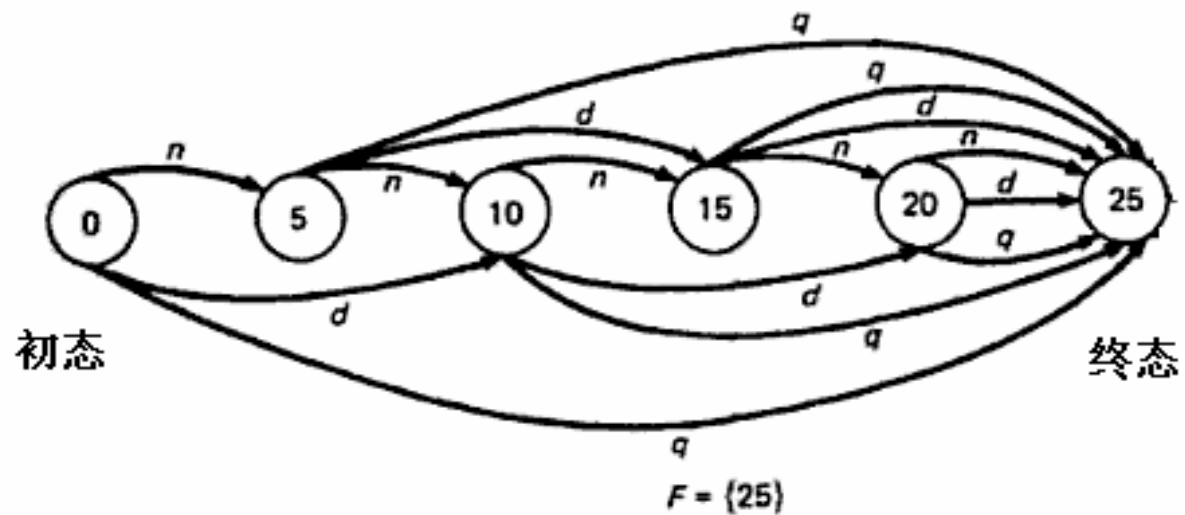
```
% {  
int user_select, total=0;  
%}  
%%  
正则表达式  { 对应的C程序动作;  
                total++;      }  
  
%%  
main()  
{  
    system("PAUSE");  
    yylex();  
    printf("共购买了%d糖果\n", total);  
}
```

程序演示……

编程作业要求（独立完成）

- 安装flex编程环境；
- 算出“自动糖果售货机（完整版）”的正则表达式（最好有作图）；
- 编写“自动糖果售货机（完整版）”的lex.l文件；
- 用flex lex.l生成lex.yy.c文件
- 用VS或VC开发环境编译并执行该程序；
- 用文本文件（或屏幕图象）记录运行过程及结果。
- 作业提交时间：2020年12月11日前。
- 提交方式：发电子邮件至: workassistant@sina.com
- 提交内容：
 - 1、“自动糖果售货机（完整版）”的lex.l文件；**不要压缩**。
 - 2、记录运行过程及结果的文本文件（或屏幕图象）。
 - 3、邮件标题为：姓名—学号—班级

完整版自动糖果售货机的DFA图



Q&A

zhangwei@cse.neu.edu.cn