



编译原理技术课程实验指导书

周汝雁 编著

上海海洋大学海洋智能信息实验教学示范中心

实验一 程序语言的重复率检查

一、实验目的

通过编写一个重复率检查程序，掌握编译器的预处理方法。

二、实验环境

不限语言，实验环境自行配置

三、实验内容

1. 打开两个 C 或 C++ 或其它程序语言文件，并显示两个文件内容；
2. 对比两个程序语言源文件，进行重复率检查，给出重复率；注：重复率:连续有 n 个词相同则计一次重复，n 可以通过界面设置进行调整，给出程序设计过程中重复率的设定及计数公式。
3. 两个文件中重复之处用红色标出或者摘录出来，注释部分不计重复率。
4. 编程语言自行选择，界面自行设计，要求有 n 设置，有重复率显示结果。

四、实验报告要求：

1、实验原理/算法原理：

- (1) 重复率公式和计算方法：

重复率=

- (2) 算法思路：

- (3) 划分单词的规则：

2、编程所用语言：

3、程序代码：

4、实验结果（需要有运行结果的截屏图片）：

5、问题及讨论：

实验二 词法分析器 1--状态转换图方法

一、实验目的

通过编写一个只包含部分保留字和单词符号（见 C 语言子集 L）的语言扫描器（词法分析器），掌握词法分析器的方法之一——状态转换图法。

二、实验环境

不限语言，实验环境自行配置

三、实验内容

1、该词法分析器的任务如下：

- （1）滤掉源程序中的无用成分，如空格；
- （2）输出分类记号供语法分析器使用，并产生两个表格（符号表）：常数表和变量表（即标识符表），分别记录曾扫描到的变量和常量，表的数据结构和包含内容自行定义；
- （3）识别非法输入，并将其标记为“出错记号”。

2、该词法分析器的功能：

以后面段落 3 中定义的 C 语言子集的源程序作为词法分析程序的输入，即源程序，源程序段存放在文件中。词法分析器打开该文件后，自文件头开始扫描源程序字符，发现符合“单词”定义的源程序字符串时，将它翻译成固定长度的单词内部表示，并查填适当的信息表（符号表），一旦发现不符合“单词”定义的源程序字符串时，给出错误提示。经过词法分析后，源程序字符串（源程序的外部表示）被翻译成具有等长信息的单词串（源程序的内部表示），并产生两个表格：常数表和变量表（即标识符表），它们分别包含了源程序中的所有常数和所有标识符。即该词法分析器可将下面 C 语言子集 L 所定义的“单词”区分开来，识别输入的源程序代码中每个“单词”是关键字，还是常量，还是变量，对于变量还要记录变量的个数，识别不同的变量，如程序段：`real xx,y;xx=1;if(xx=1)y=5;`要能识别出有 2 个变量，分别是 `xx,y`，第 1 次扫描到 `xx`，识别出是(6,0) 第 2 次、第 3 次扫描到 `xx`，识别出仍是(6,0)；第 1 次扫描到 `y`，识别出是(6,1)，第 2 次扫描到 `y`，识别出仍是(6,1)；对于常数采用同样的识别方法。

3、C 语言子集 L---保留字不区分大小写

$L = \{ \text{IF, THEN, ELSE, INT, CHAR, FOR, =, >=, ==, +, /, \%, ++, ", ;} \cup \{ \text{常数} \} \cup \{ \text{变量, 即标识符} \}$ ，变量只由字母构成，字母个数小于 10。

4、实验要求：

- （1）词法分析器构造原理，单词符号（记号）的分类编码表，状态转换图设计绘制写在实验报告中；
- （2）进行词法分析器的测试：测试例程（一小段程序）、测试结果与测试结果分析。

5、例子:

本例中单词符号（记号）的种类:

- 1、保留字（关键字）；
- 2、分隔符；
- 3、运算符；
- 4、常数；
- 5、标识符

（单词符号的分类可以自己规定，请在实验报告中给出分类编码表）

测试用例

第一段:

```
for (i=1;i<=100;i++)
{   printf("%d ", i );
}
```

over

运行词法分析程序后，显示结果如下:

```
for   (for,16)
(     (,err)
i     (i,51)
=     (=,31)
1     (1,41)
:     (: ,22)
i     (i ,51)
<=   (<=,err)
100  (100,42)
:     (: ,22)
i     (i ,51)
++   (++,37)
)     (,err)
{     ({,err)
Printf (Printf,52)
(     (,err)
"     (" ,21)
%     (% ,36)
d     (d,53)
"     (" ,21)
,     (,err)
i     (i,51)
)     (,err)
:     (: ,22)
}     (},err)
```

over

常数表中的内容为: 1, 100

变量表（标识符表）中的内容为：i, printf, d

更多测试用例

第二段：

```
{ int n,i,kk;
  printf("n=?"); scanf("%d",&n);
  for (i=2; i<=kk; i++)
    if(n%i==0) break;
  if(i<n) printf("%d no\n",n);
  else printf("%d yes\n",n);
  return 0;
}
```

over

第三段：

```
char c;
int letters=0,space=0,digit=0,other=0;
printf("\n");
while((c=getchar())!='\n')
{
  if (( c>='a' && c<='z' ) || ( c>='A' && c<='Z' ) ) letters++;
  else if (c==' ') space++;
  else if (c>='0' && c<='9') digit++;
  else other++;
}
printf(letters,space,digit,other);
over
```

四、实验报告要求：

1、实验原理/算法原理：

包括本实验词法分析器构造原理，单词符号（记号）的分类编码表，状态转换图设计

2、编程所用语言：

3、程序代码：

4、实验结果（需要有运行结果的截屏图片）：

5、问题及讨论：

实验三 词法分析器 2—自动机方法

一、实验目的

通过编写一个扫描器（词法分析器），掌握词法分析器的构造方法之一：自动机方法（模拟 DFA 法）。

二、实验环境

不限语言，实验环境自行配置

三、实验内容

1.（选做 a、b、c、d 中的一个，但实验报告要对没有选择的另外三个如何做分别进行说明，给出思路，实验分数分配：选做 d 最高，c 居中，a 和 b 分数相同，）

- a. 采用模拟 DFA 算法编写一个扫描器（词法分析器），用来识别：由任意个 a 或 b 开始（任意 a、b 串开始），后接 bb，再自加或自减 1 的字符串，即正规式 $r=(a|b)^*bb(+|-)1$ 描述的语言 $L(r)$ 。
- b. 采用模拟 DFA 算法编写信箱识别器，用来识别输入串是否为 126 或 163 信箱，若输入串是 126 或 163 信箱，输出“信箱地址正确”，若输入串非 126 或 163 信箱，则输出“错误信箱地址”。信箱地址简化为由任意个 a 或 b 开始（任意 a、b 串开始），后接@，再接 126，或接 163，再接 m 的字符串，请先写出正规式 r，再编写信箱识别器。
- c. 采用模拟 DFA 算法编写一个扫描器（词法分析器），用来识别：由任意个 a 或 b 开始（任意 a、b 串开始），后接 bb，再接 C 语言的+、-、*、/、%这 5 个运算符，再接 b 的表达式，例如 a+b, ab*b, aa/b 都不可以识别，abb+b, abbbb*b, aabb/b 都可以识别。正规式 r 自行写出。
- d. 采用模拟 DFA 算法编写一个扫描器（词法分析器），用来识别：由任意个 a 或 b 开始（任意 a、b 串开始），后接 bb，再接 C 语言关系运算，再接 1 的表达式，例如 a>=1, ab==1, aa>1, b<1 都不可以识别，abb>=1, abbbb==1, aabb>1, bb<1 都可以识别。正规式 r 自行写出。

2. 实验要求:

(1) 该词法分析器的任务:

滤掉源程序中的无用成分，如空格； 对于多行输入以;为每个句子的分隔符；

(2) 该词法分析器的功能:

从键盘读入或打开文件读入一行或多行字符串，词法分析器读入字符串后扫描源字符串，若发现该行符合正规式 r 描述的字符串时，输出“yes”或“可接受”或“可识别”或“合法”，否则输出“no”或“不可接受”或“不可识别”或“不合法”。

测试样例：

实验选项 a

测试清单 1:

输入：

aa+1:

ba-11:

babbaa-1:

输出显示结果如下：

aa+1 不可识别

ba-11 不可识别

babbaa-1 不可识别

测试清单 2:

输入：

bb-1:

abb-22:

babbbb+1:

输出显示结果如下：

bb-1 合法

abb-22 非法

babbbb+1 合法

实验选项 b

测试清单 1:

输入：

aa@126m:

b@163m:

bbb@126m1:

输出显示结果如下：

aa@126m 合法

b@163m 合法

bbb@126m1 非法

其它选项请自行设计测试清单

四、实验报告要求：

1、实验原理/算法原理（本词法分析器构造过程）：

2、编程所用语言:

3、程序代码:

4、实验结果（需要有运行结果的截屏图片）:

5、问题及讨论:

实验四 语法分析器 1--预测分析器构造

一、实验目的

通过编写一个预测分析器,掌握自上而下构造语法分析器的方法之一-----预测分析法。

二、实验环境

不限语言,实验环境自行配置

三、实验内容

1、该语法分析器的任务如下:

(1) 可识别程序代码中符合以下文法中的 3-5 个,返回该语句为何种语句,或提示出错

文法 G1:

```
L → E
E → E = T | T
E → E + T | T
T → T - F | F
F → (E) | id
```

文法 G2:

```
L → E
E → E = T | T
E → E * T | T
T → T / F | T % F | F
F → (E) | id
```

文法 G3:

```
D → TL
T → int | flort
L → id R
R → , id R | ε
```

文法 G4:

```
L → E; L | ε
E → TE'
E' → +TE' | -TE' | ε
T → FT'
T' → *FT' | /FT' | mod FT' | ε
F → (E) | id | num
```

文法 G5: (P49 例 3.7, P58 例 3.15 的文法)

```
stmt → if expr then stmt else stmt
      | if expr then stmt
```

|other

(2) 出错处理可自行设计,并添加在预测分析表中,参考教材 P59-62。

2、实验要求:

测试例子如下蓝色部分,或其它程序中的某几行语句、测试结果与测试结果分析。

```
#include <stdio.h>
#include <math.h>
int main ( )
{int a,b,c,disc,x1,x2,p,q;
scanf("%lf%lf%lf",&a,&b,&c);
sqrt=(b*b)/(4*a*c)
disc=b*b-a*c;
if (disc<0) p=2.0*a;
else p=-b/(2.0*a)
printf("x1=%7.2f\nx2=%7.2f\n",x1,x2);
}
return 0;
}
```

四、实验报告要求:

1、实验原理/算法原理:

语法分析器构造原理,以及构造过程需要用到的集合、表格等

2、编程所用语言:

3、程序代码:

4、实验结果(需要有运行结果的截屏图片):

5、问题及讨论:

实验五 语义分析和语法制导翻译

一、实验目的

通过编写一个语义分析器，掌握语义分析的方法—语法制导翻译。

二、实验环境

不限语言，实验环境自行配置

三、实验内容

1.该语义分析器的任务如下：

(1) 在实验 4 的基础上完成声明语句和赋值语句（语法为 G1 或 G2）的语法制导翻译，给出符号表并为变量分配空间（地址空间自行设计起始位置，起始地址显示在符号表中即可）或三元式中间代码

文法 G1:

```
L → E
E → E = T | T
E → E + T | T
T → T - F | F
F → (E) | id
```

文法 G2:

```
L → E
E → E = T | T
E → E * T | T
T → T / F | T % F | F
F → (E) | id
```

文法 G3:

```
D → TL
T → int | flort
L → id R
R → , id R | ε
```

文法 G4:

```
L → E ; L | ε
E → TE'
E' → +TE' | -TE' | ε
T → FT'
T' → *FT' | /FT' | mod FT' | ε
F → (E) | id | num
```

2. 测试例子如下蓝色部分，或其它程序中的某几行语句、测试结果与测试结果分析。

```
#include <stdio.h>
#include <math.h>
int main ( )
{int a,b,c,disc,x1,x2,p,q;
scanf("%lf%lf%lf",&a,&b,&c);
sqrt=b*b/4*a*c
disc=b+b-a+c;
if (disc<0) p=2.0*a;
else p=-b/(2.0*a)
printf("x1=%7.2f\nx2=%7.2f\n",x1,x2);
}
return 0;
```

四、实验报告要求:

1、实验原理/算法原理:

语义规则和翻译方案

2、编程所用语言:

3、程序代码:

4、实验结果（需要有运行结果的截屏图片）:

5、问题及讨论:

实验六 人机对话

一、实验目的

通过编写一个人机对话程序，了解语言的理解和人工智能相关知识、技术和方法。

二、实验环境

不限语言，实验环境自行配置

三、实验内容

运行该程序可进行人与机器的对话，即：计算机能根据输入的文字（英语或中文）进行相应的对话。

四、实验报告要求：

1、实验原理/算法原理：

给出程序的基本设计思路、语言判断方法和回答策略。

2、编程所用语言：

3、程序代码：

4、实验结果（需要有运行结果的截屏图片）：

5、问题及讨论：



实验报告

题目：_____

学院：信息学院

专业：

班级：

学号：

姓名：

年 月 日