



华南师范大学

本科学生实验（实践）报告

院 系：计 算 机 学 院

实验课程：编译原理

实验项目：XLEX-词法自动生成器（测试报告）

指导老师：黄煜廉

开课时间：2023 ~ 2024 年度第 1 学期

专 业：计算机科学与技术

班 级：计科 1 班

学 生：李达良

学 号：20203231004

华南师范大学教务处

华南师范大学实验报告

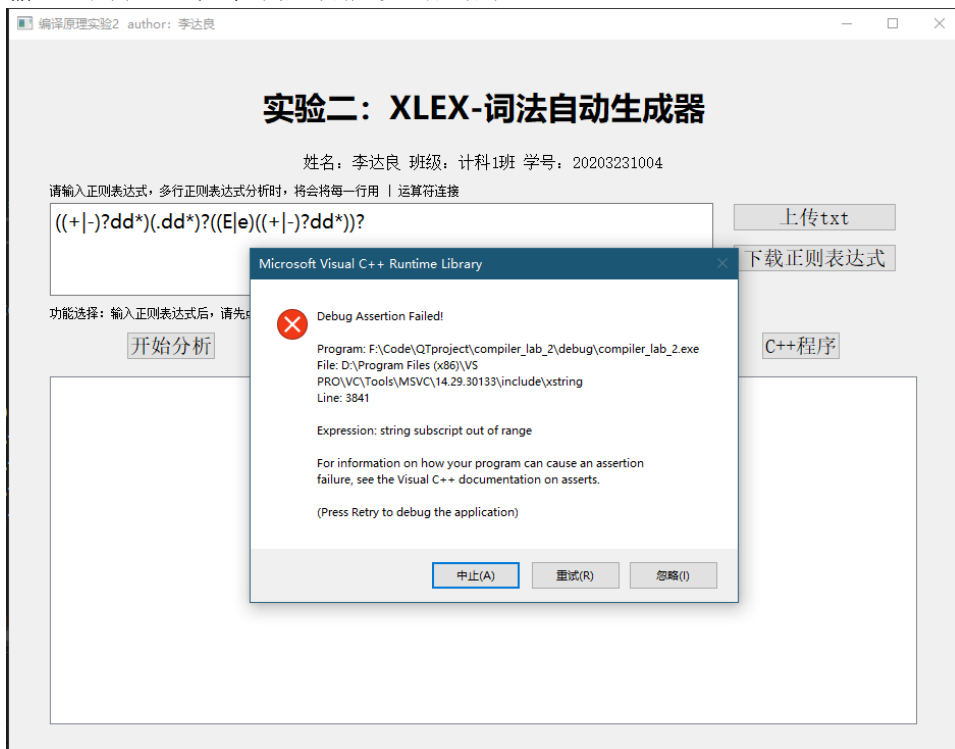
学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分 _____

一、测试样例

$((+|-)?dd^*)(.dd^*)((E|e)((+|-)?dd^*))?$

二、测试内容

输入正则表达式，检测是否能够生成结果



错误原因：

1. 在题目中，+号为正闭包，不应该出现在正则表达式当作字符，应该用字母代替

三、选做实验要求：

(1) 扩充正则表达式的运算符，如 $[]$ 、正闭包 $(+)$ 等。

2. 在程序中有 isChar 检测，如果不是字母，不作为普通字符：

```
/*=====正则表达式前置处理=====*/  
  
// 判断是不是字符  
bool isChar(char c)  
{  
    if ((c >= 'a' && c <= 'z') || (c >= 'A' && c <= 'Z'))  
        return true;  
    return false;  
}
```

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分

因此，需要将正则表达式中的非字母字符换成字母字符，方可进行分析：

+替换成 a

-替换成 b

.替换成 c

得到正则表达式为：((a|b)?dd*)(cdd*)?((E|e)((a|b)?dd*))?

再次进行测试：

得到的 NFA：

编译原理实验2 author: 李达良

实验二：XLEX-词法自动生成器

姓名：李达良 班级：计科1班 学号：20203231004

请输入正则表达式，多行正则表达式分析时，将会将每一行用 | 运算符连接

((a|b)?dd*)(cdd*)?((E|e)((a|b)?dd*))?

上传txt

下载正则表达式

功能选择：输入正则表达式后，请先点击开始分析，再点击其他按钮查看结果

开始分析

NFA

DFA

DFA最小化

C++程序

	标志	ID	#	E	a	b	c	d	e
1	-	6	4,7						
2		4	0,2						
3		7	8						
4		8						9	
5		9	12						
6		12	10,13						
7		10						11	
8		13	22						
9		22	14,23						
10		14					15		
11		23	11						

得到的 DFA：

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分

编译原理实验2 author: 李达良

实验二：XLEX-词法自动生成器

姓名：李达良 班级：计科1班 学号：20203231004

请输入正则表达式，多行正则表达式分析时，将会将每一行用 | 运算符连接

功能选择：输入正则表达式后，请先点击开始分析，再点击其他按钮查看结果

标志	状态集合	E	a	b	c	d	e
5	{25,29,30,32,3...}		{31,35,37,38}	{33,35,37,38}		{39,40,42,43,45}	
6	{15,16}					{17,18,20,21,2...}	
7 +	{10,11,13,14,2...}	{25,29,30,32,3...}			{15,16}	{10,11,13,14,2...}	{27,29,30,32,3...}
8	{27,29,30,32,3...}		{31,35,37,38}	{33,35,37,38}		{39,40,42,43,45}	
9	{31,35,37,38}					{39,40,42,43,45}	
10	{33,35,37,38}					{39,40,42,43,45}	
11 +	{39,40,42,43,45}					{40,41,43,45}	
12 +	{17,18,20,21,2...}	{25,29,30,32,3...}				{18,19,21,23,2...}	{27,29,30,32,3...}
13 +	{40,41,43,45}					{40,41,43,45}	
14 +	{18,19,21,23,2...}	{25,29,30,32,3...}				{18,19,21,23,2...}	{27,29,30,32,3...}

得到的 DFA 最小化：

编译原理实验2 author: 李达良

实验二：XLEX-词法自动生成器

姓名：李达良 班级：计科1班 学号：20203231004

请输入正则表达式，多行正则表达式分析时，将会将每一行用 | 运算符连接

功能选择：输入正则表达式后，请先点击开始分析，再点击其他按钮查看结果

标志	状态集合	E	a	b	c	d	e
1 -	0		2	2		1	
2 +	1	5			6	1	5
3	2					1	
4 +	3					3	
5 +	4	5				4	5
6	5		7	7		3	
7	6					4	
8	7					3	

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004

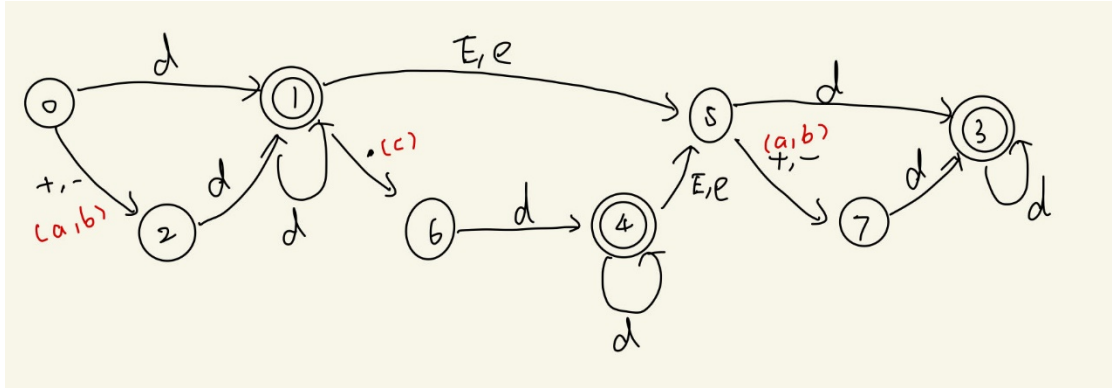
专 业 计算机科学与技术 年级、班级 2021 级 1 班

课程名称 编译原理 实验项目 XLEX-词法自动生成器 (测试报告)

实验时间 2023 年 11 月 7 日

实验指导老师 黄煜廉 实验评分

通过手画作图，可以得到 DFA 最小化标准答案为：



对比程序生成的答案是**完全正确**的。

生成的代码（已将 abc 转换为对应符号）：

```
#include <iostream>
#include <string>

using namespace std;

int main() {
    string input;
    cout << "Enter input string: ";
    cin >> input;
    int currentState = 0;
    int length = input.length();
    for (int i = 0; i < length; i++) {
        char c = input[i];
        switch (currentState) {
            case 0:
                switch (c) {
                    case 'E':
                        cout << "Error: Invalid input character '" << c << "' " << endl;
                        break;
                    case '+':
                        currentState = 2;
                        break;
                    case '-':
```

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004

专 业 计算机科学与技术 年级、班级 2021 级 1 班

课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）

实验时间 2023 年 11 月 7 日

实验指导老师 黄煜廉 实验评分

```
        currentState = 2;
        break;
    case '.':
        cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;
        break;
    case 'd':
        currentState = 1;
        break;
    case 'e':
        cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;
        break;
    default:
        cout << "Error: Invalid input character '" << c << "'" << endl;
        return 1;
    }
    break;
case 1:
    switch (c) {
    case 'E':
        currentState = 5;
        break;
    case '+':
        cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;
        break;
    case '-':
        cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;
        break;
    case '.':
        currentState = 6;
        break;
    case 'd':
        currentState = 1;
        break;
    case 'e':
```

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分

```
        currentState = 5;
        break;
    default:
        cout << "Error: Invalid input character '" << c << "'" << endl;
        return 1;
    }
    break;
case 2:
    switch (c) {
        case 'E':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case '+':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case '-':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case '.':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case 'd':
            currentState = 1;
            break;
        case 'e':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
    default:
        cout << "Error: Invalid input character '" << c << "'" << endl;
        return 1;
    }
    break;
case 3:
```

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分 _____

```
switch (c) {  
case 'E':  
    cout << "Error: Invalid input character '" << c << "'" << endl;  
return 1;  
    break;  
case '+':  
    cout << "Error: Invalid input character '" << c << "'" << endl;  
return 1;  
    break;  
case '-':  
    cout << "Error: Invalid input character '" << c << "'" << endl;  
return 1;  
    break;  
case '.':  
    cout << "Error: Invalid input character '" << c << "'" << endl;  
return 1;  
    break;  
case 'd':  
    currentState = 3;  
    break;  
case 'e':  
    cout << "Error: Invalid input character '" << c << "'" << endl;  
return 1;  
    break;  
default:  
    cout << "Error: Invalid input character '" << c << "'" << endl;  
    return 1;  
}  
break;  
case 4:  
    switch (c) {  
case 'E':  
        currentState = 5;  
        break;  
case '+':  
        cout << "Error: Invalid input character '" << c << "'" << endl;  
return 1;  
        break;
```


华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分

```
        case '-':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

        break;
        case '.':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

        break;
        case 'd':
            currentState = 4;
            break;
        case 'e':
            currentState = 5;
            break;
        default:
            cout << "Error: Invalid input character '" << c << "'" << endl;
            return 1;
    }
    break;
case 5:
    switch (c) {
        case 'E':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case '+':
            currentState = 7;
            break;
        case '-':
            currentState = 7;
            break;
        case '.':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case 'd':
            currentState = 3;
            break;
```

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分

```
        case 'e':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        default:
            cout << "Error: Invalid input character '" << c << "'" << endl;
            return 1;
    }
    break;
case 6:
    switch (c) {
        case 'E':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case '+':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case '-':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case '.':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        case 'd':
            currentState = 4;
            break;
        case 'e':
            cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

            break;
        default:
            cout << "Error: Invalid input character '" << c << "'" << endl;
            return 1;
    }
}
```

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分 _____

```
        break;
    case 7:
        switch (c) {
            case 'E':
                cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

                break;
            case '+':
                cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

                break;
            case '-':
                cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

                break;
            case '.':
                cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

                break;
            case 'd':
                currentState = 3;
                break;
            case 'e':
                cout << "Error: Invalid input character '" << c << "'" << endl;
return 1;

                break;
            default:
                cout << "Error: Invalid input character '" << c << "'" << endl;
                return 1;
        }
        break;
    }
}

switch (currentState) {
case 1:
    cout << "Accepted" << endl;
    break;
case 3:
```

华南师范大学实验报告

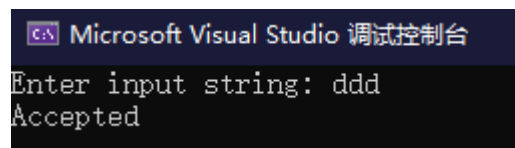
学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分

```
        cout << "Accepted" << endl;  
        break;  
case 4:  
        cout << "Accepted" << endl;  
        break;  
default:  
        cout << "Not Accepted" << endl;  
}  
return 0;  
}
```

对程序进行测试：

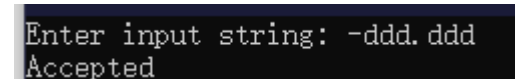
合法部分：

① ddd



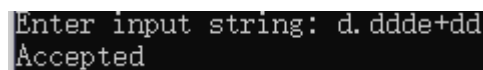
```
Microsoft Visual Studio 调试控制台  
Enter input string: ddd  
Accepted
```

② -ddd.ddd



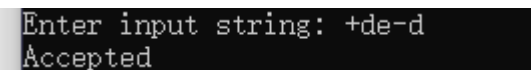
```
Enter input string: -ddd.ddd  
Accepted
```

③ d.ddde+dd



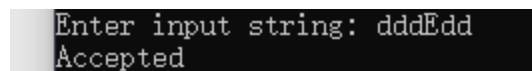
```
Enter input string: d.ddde+dd  
Accepted
```

④ +de-d



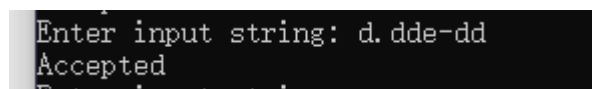
```
Enter input string: +de-d  
Accepted
```

⑤ dddEdd



```
Enter input string: dddEdd  
Accepted
```

⑥ d.dde-dd



```
Enter input string: d.dde-dd  
Accepted
```

非法：

⑦ -.

华南师范大学实验报告

学生姓名 李达良 学 号 20203231004
专 业 计算机科学与技术 年级、班级 2021 级 1 班
课程名称 编译原理 实验项目 XLEX-词法自动生成器（测试报告）
实验时间 2023 年 11 月 7 日
实验指导老师 黄煜廉 实验评分

```
Enter input string: -.
Error: Invalid input character '.'
```

⑧ +Ed

```
Microsoft Visual Studio 调试控制台
Enter input string: +Ed
Error: Invalid input character 'E'
```

⑨ -d.dE-df

```
Microsoft Visual Studio 调试控制台
Enter input string: -d.dE-df
Error: Invalid input character 'f'
```

⑩ d.dE+

```
F:\Code\Project1\Project1\x64\Debug\Project1.exe
Enter input string: d.dE+
Not Accepted
```

三、小结

本次样例由于程序中带有正负号与正闭包检测冲突，同时程序仅支持字母作为单个字符输入，在测试时会直接报错，因此样例测试不通过。

但是，将正则表达式中的非字母字符换成字母字符后，再进行分析后，与手工做法得到的答案完全吻合，因此处理后的正则表达式样例是**通过**的。

最后生成的 C++ 代码，通过 10 个测试样例证明生成的代码也是符合正则表达式规则的，因此本次实验二**测试通过**。