**软件课程设计报告**

**题目：词法分析与语法分析**

**（JAVA版）**

**班级：11065501**

**学号：1116120206**

**姓名：王丹**

**2014/6/26**

**目录**

**一、概述**

**1.课程设计的背景**

**2.课程设计的目的**

**3、课程设计的要求**

**二、总体设计方案**

**1、设计思路**

**2.系统总体UML（用例图和类图）**

**三、详细设计**

**1、C++/Java子模块（状态图，活动图、数据结构、运行结果）**

**2、DFA子模块（状态图，活动图、数据结构、运行结果）**

**3、LL1子模块（状态图，活动图、数据结构、运行结果）**

**4、OPG子模块（状态图，活动图、数据结构、运行结果）**

**四、疑难解答**

**1、**

**2.**

**3.**

**五、课程设计总结**

**一、概述**

**1、课程设计背景**

一个编译程序的工作过程一般可以划分为五个阶段：词法分析、语法分析、语义分析与中间代码生成、优化、目标代码生成。每个阶段都是从上一个阶段得到结果，对他进行分析，并且根据一些外部环境(例如符号表等)得到最终的输出结果。要构造一个编译程序，可以按照这样的阶段来分别构造，最后来连调。

编译程序是一极其庞大而又复杂的系统，掌握它比较苦难。但是一旦对其掌握，对以后的程序语言设计，系统软件分析，系统软件设计，形式语言研究等方面都是非常有好处的。

**2、课程设计的目的**

掌握计算机语言的词法分析程序的开发方法以及掌握计算机语言的语法分析程序设计与属性文法应用的实现方法。通过自己动手实现上课讲述的编译的相关程序，加深对所学知识的理解，进而培养“计算思维能力”——模型化、抽象思维能力、逻辑思维能力以及算法设计与分析能力，程序设计与实现能力和计算机系统的认知、分析、设计和应用能力 。

**3、课程设计的要求**

（1）给出C++/Java的词法和语法定义，构造一个编译程序，它主要可以完成如下功能：

1、读入某个已经编辑好的C++/Java源程序文件，通过词法分析器，生成token串，同时检查词法错误

2、语法分析器将产生的token串作为输入，进行语法分析，同时检查语法错误；

（2）以文法作为输入，分别测试词法和语法。

词法分析部分要求： 1． 根据所给的正规文法，画出NFA，再化简为DFA。

语法分析部分要求：

考虑LL1要求：1.求所给文法的First集、Follow集和Select集。 2. 利用Select判断文法是否是LL1文法。3． 若是LL1文法，给出预测分析表。4.利用预测分析表，判断给定字符串是否是该文法是句子。

考虑OPG要求：1.求所给文法的FirstVT集和LastVT集。 2.判断所给文法是否是OPG文法。3． 若是，给出算符优先分析表。4.利用算符优先分析表，判断给定字符串是否是该文法是句子。

**二、总体设计方案**

**1、设计思路**

C++/Java词法分析的功能是从源程序中读入一个个字符，一句一定的构词规则，识别出各类有用的单词。

程序首先定义一个数组，存放所有界符。当程序读入的一个字符时，首先判断该字符是否属于界符或空格。若是，则判断临时字符串所属种别。否则，继续读入下一字符。直到程序结束。

DFA词法分析的功能是根据所给文法构造其NFA，并从初态开始，走至终态，以数字作为新状态名，构造DFA。

LL1分析的功能是根据文法构造First图，然后根据First图寻找每一非终结符的First集，Follow集类似。实现该功能主要应用递归函数。

OPG分析的功能是根据文法构造FirstVT图，然后根据FirstVT图寻找每一非终结符的FirstVT集，LastVT集类似。实现该功能主要应用递归函数。

字符串判断是根据DFA、LL1的预测分析表或OPG的算符优先分析表，按照各自的规则进行移进规约匹配至成功或失败。

**2.UML**

系统整体用例图



系统整体类图



子模块UML图详见详细设计。

**三、详细设计**

**1、C++/Java子模块**

本模块将C++/Java源程序分解成token串，以（序号；类型；字符串；行号）方式输出至屏幕上。然后根据token串判断语法错误。

**（1）状态图**

**（2）活动图**

**（3）数据结构**

Cifa\_CPP属性表（只列出主要的）

|  |  |  |
| --- | --- | --- |
| 属性 | 类型 | 作用 |
| CORJ | boolean |  |
| boundedStr | String[] |  |
| currentStr | String |  |
| keyWordStr | String[] |  |
| keyWords | String[] |  |
| number | int |  |
| row | int |  |
| tokenList | List<Token> |  |
| type | Int[] |  |

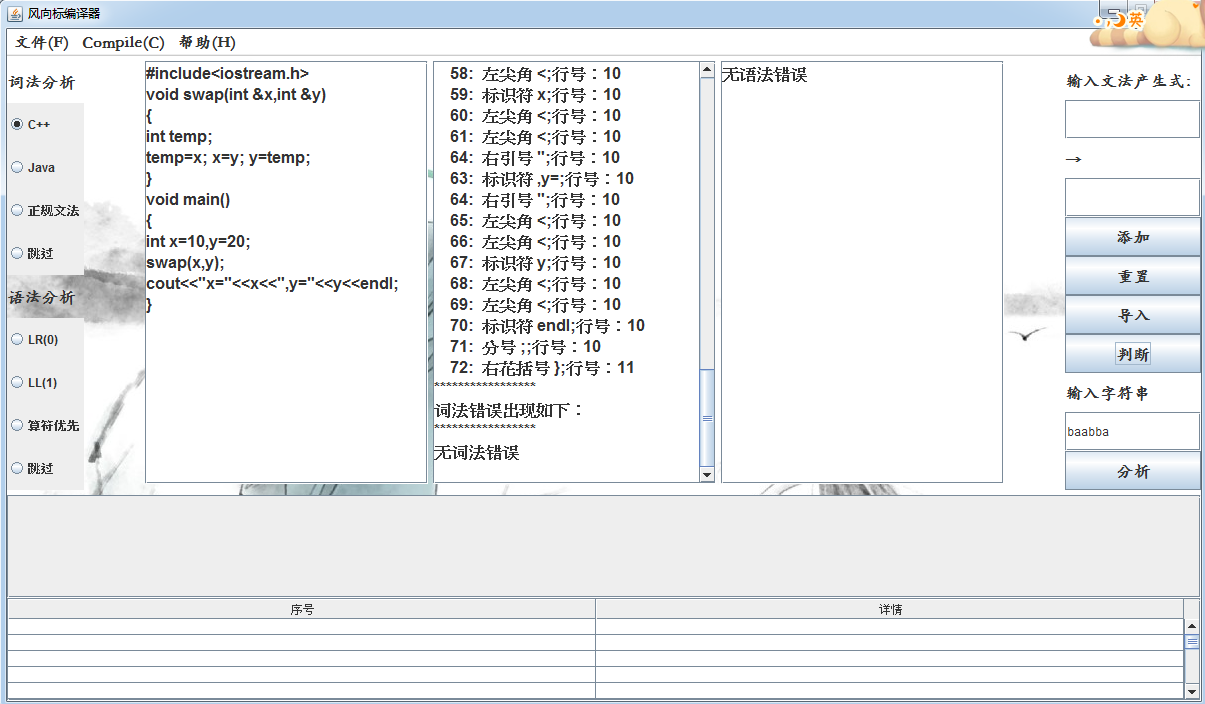
Token

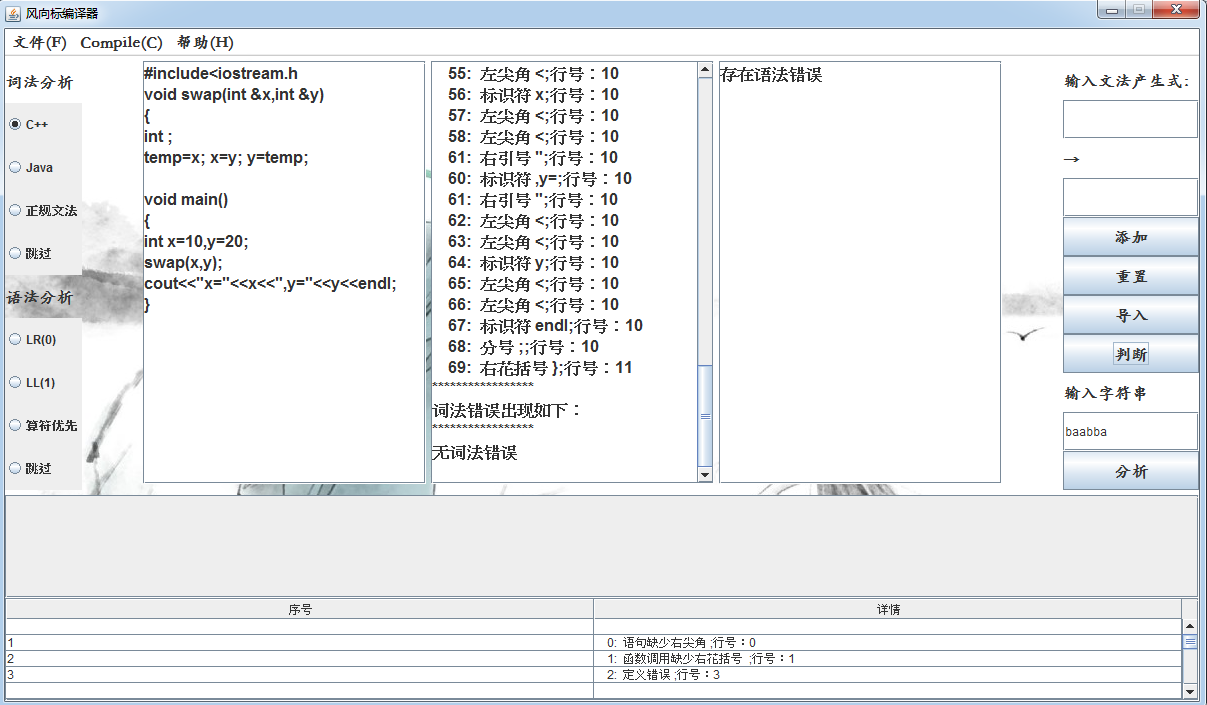
Yufa\_CPP

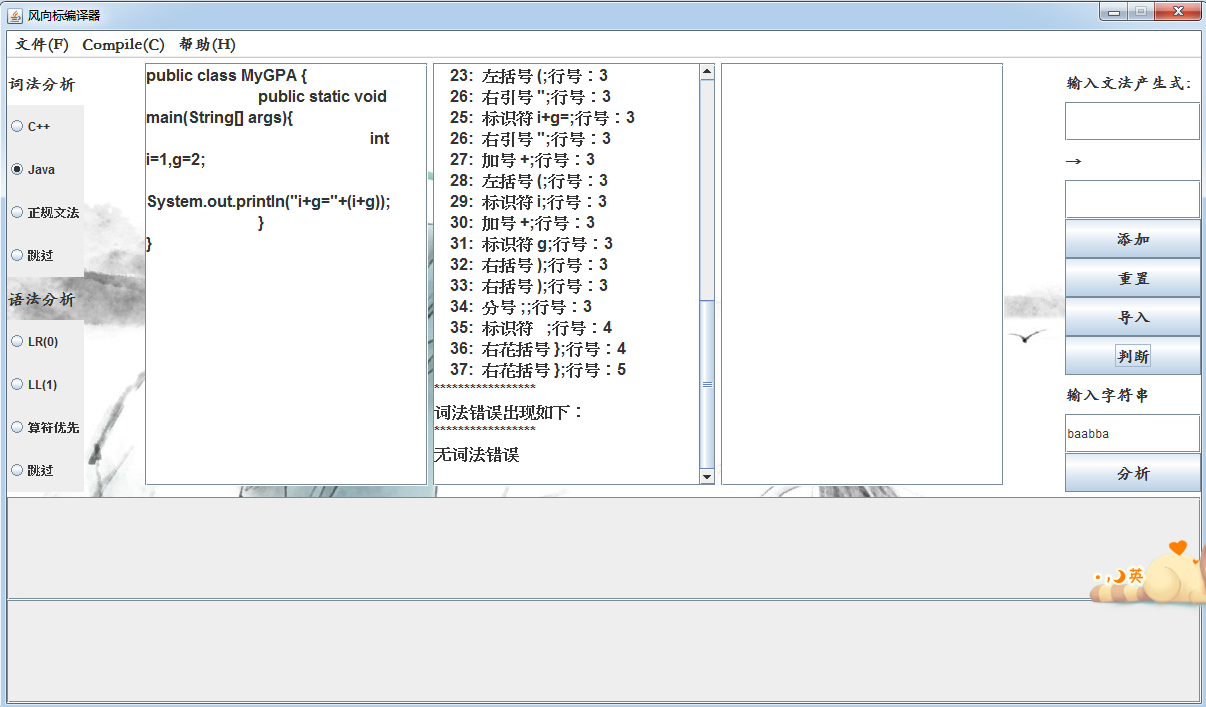
|  |  |  |
| --- | --- | --- |
| 属性/方法名 | 类型/返回值 | 作用 |
| currentInt | int |  |
| currentRow | int |  |
| currentStr | int |  |
| errorInt | int |  |
| errorList | List<ErrorInfo> |  |
| identifierList | List<Token> |  |
| row |  |  |
|  |  |  |

ErrorInfo

**（4）运行结果**

****

****

****

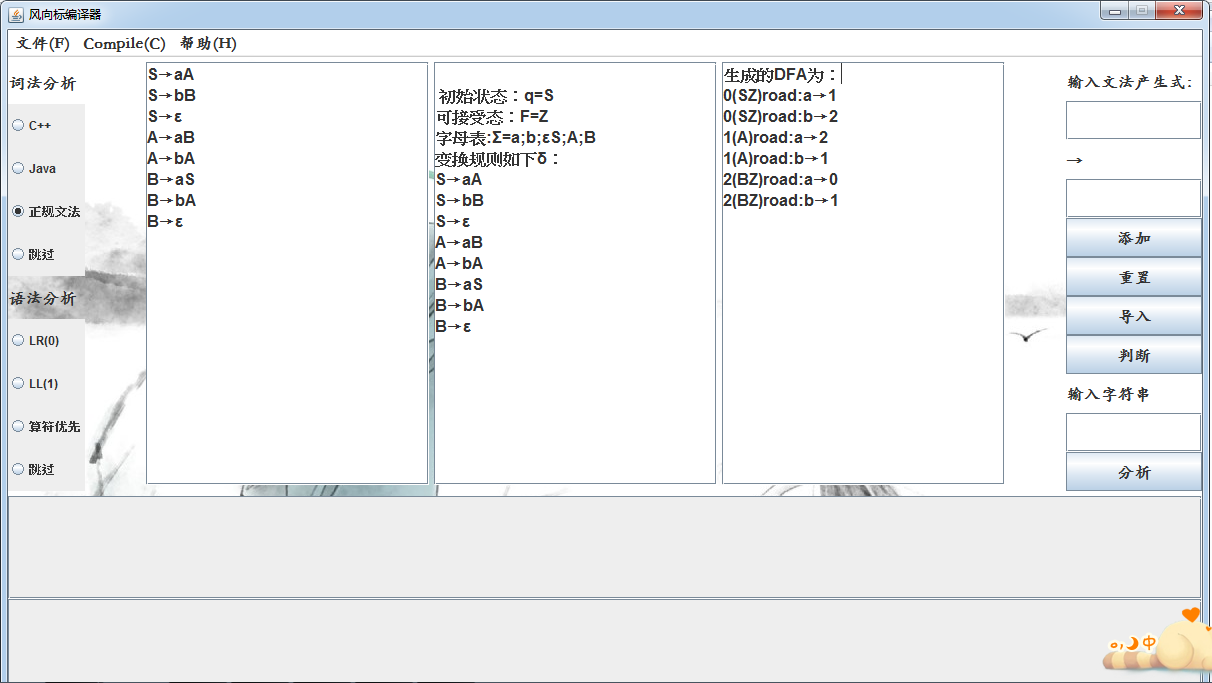
**2、DFA子模块**

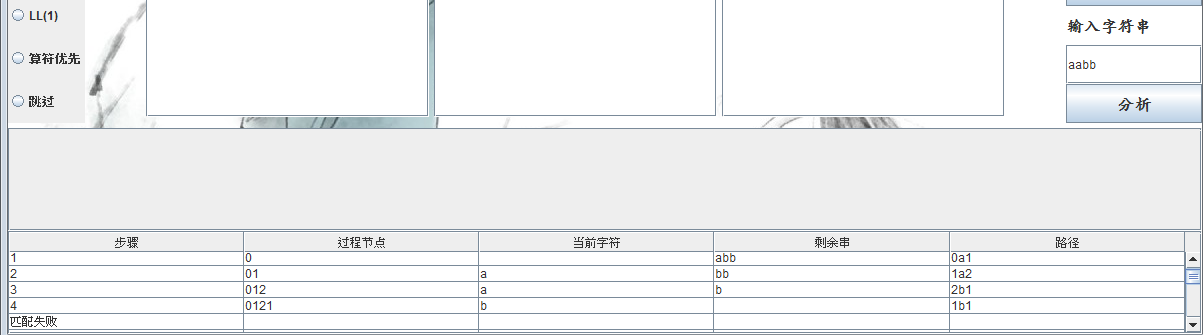
**（1）状态图**

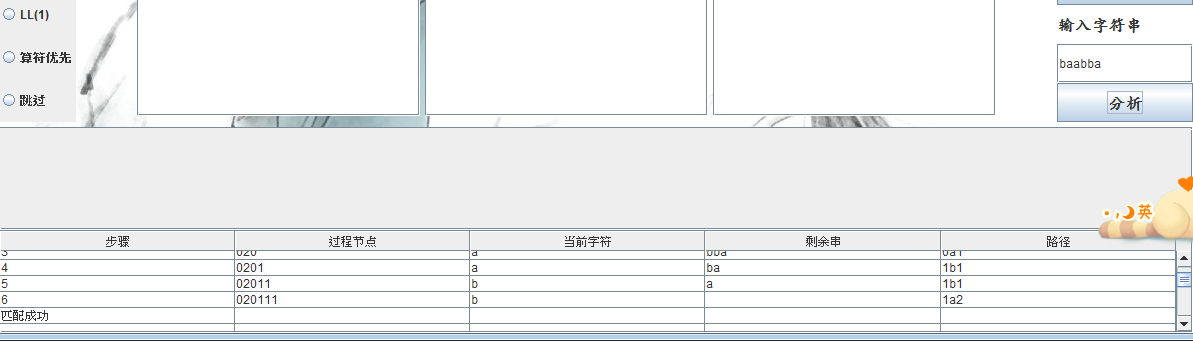
**（2）活动图**

**（3）数据结构**

**（4）运行结果**

****

****

****

**3、LL1子模块**

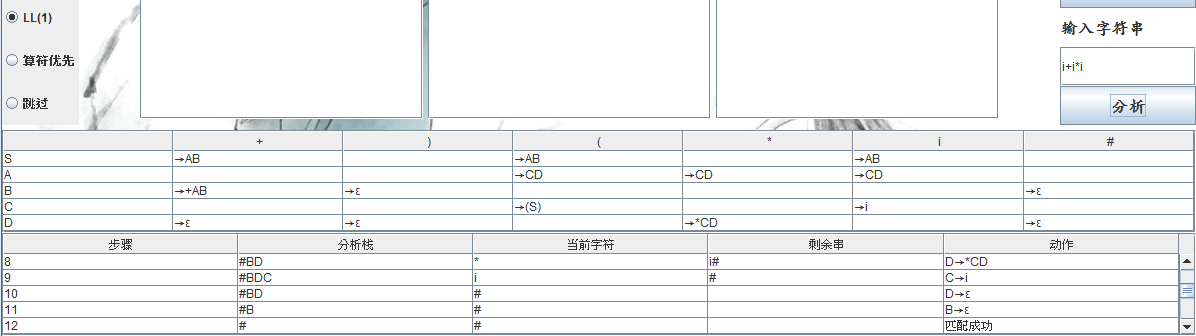
**（1）状态图**

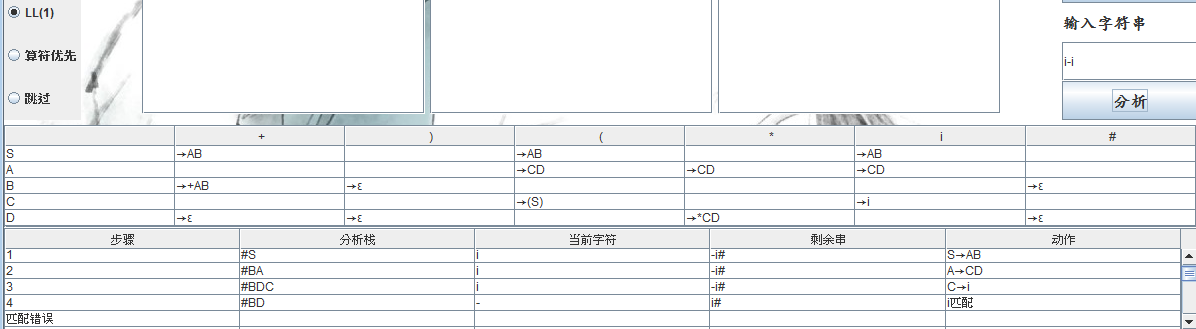
**（2）活动图**

**（3）数据结构**

**（4）运行结果**

****

****

****

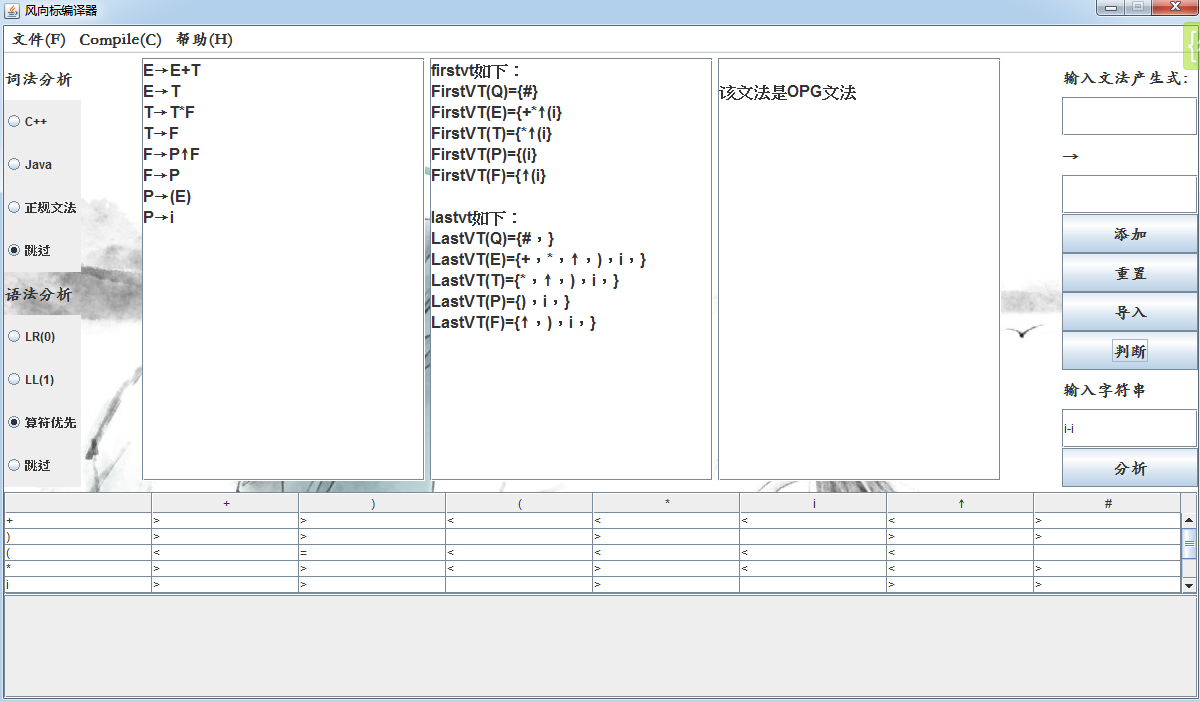
**4、OPG子模块**

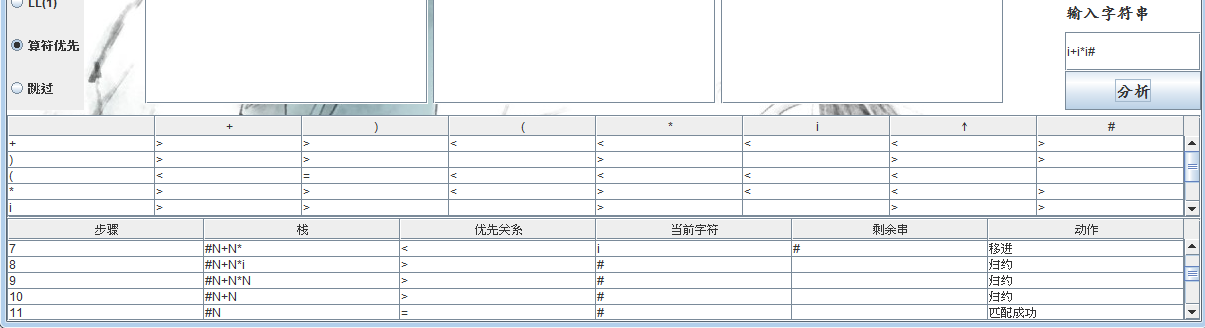
**（1）状态图**

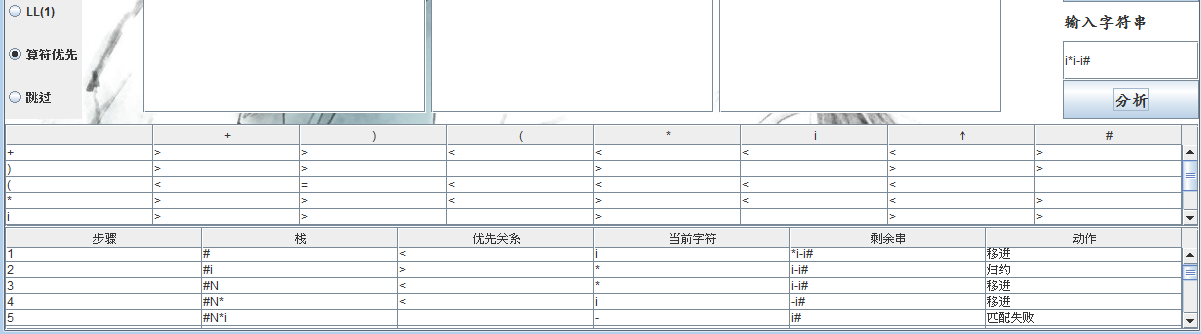
**（2）活动图**

**（3）数据结构**

**（4）运行结果**

****

****

****

**四、疑难解答**

**1、**

**2.**

**3.**

**五、课程设计总结**