



验收文档

唐子钰 2020212271

环境

visual-studio 2019, c++14 标准版, windows10

词法分析

实现内容

我基于 c++ 实现了 c 语言的词法分析程序，完成了包括但不限于以下内容

- 实现了输入字符的**双 buffer 缓冲区**，支持每次读入一个字符或读到第一个非分割符，包括 `get_char()`、`read_buf()`、`retract()`
- **识别并存储 c 源程序每个 token**，统计时能以记号 **<token_name,value> line:number** 的形式输出
- 识别并跳过注释，识别双引号字符串或单引号常字符
- 统计源程序语句行数、各类单词总数、字符总数等并在统计时输出统计结果
- 设计了几类**词法错误**并能报告错误信息包括错误所在行
- 支持**完整一次扫描**对源程序进行词法分析（跳过错误），也支持每次**单独识别一个 token**

C++

```
1  /* 全局进行一次词法分析flag=1只识别一个token */
2  void Lexical_Analysis::analyze(int flag) {
3      char c;
4      while (1) {
5          switch (state)
6          {
7              case 0://识别新词
8                  memset(token, 0, sizeof(token));
9                  token_len = 0;
10                 c = get_char();
11             case 1:// 识别标识符
12             case 2-5: // 识别符号数和非符号数
13             case 6:// 算术运算符;注意如果是//或者/*转到注释对应的状态
14             case 7-8:// 关系运算符
15             case 9:// 位运算符
16             case 10:// separator分隔符
17             case 11-12:// 双引号字符串-支持换行(单\ )和转义
18             case 13:// 单引号字符串，只允许一个字符且不允许空串且\不hint换行
19             case 14:// 双斜杠注释,\可以hint转义或者换行
20             case 15:// /**/段落注释
21             case 101: {//[Error] unable to find numeric literal operator 'operator'"a'
22             default:
23                 break;
24             }
25         }
26     }
```

词法分析主程序框架

测试

`lexical_ana()` 词法分析进行完整一次扫描

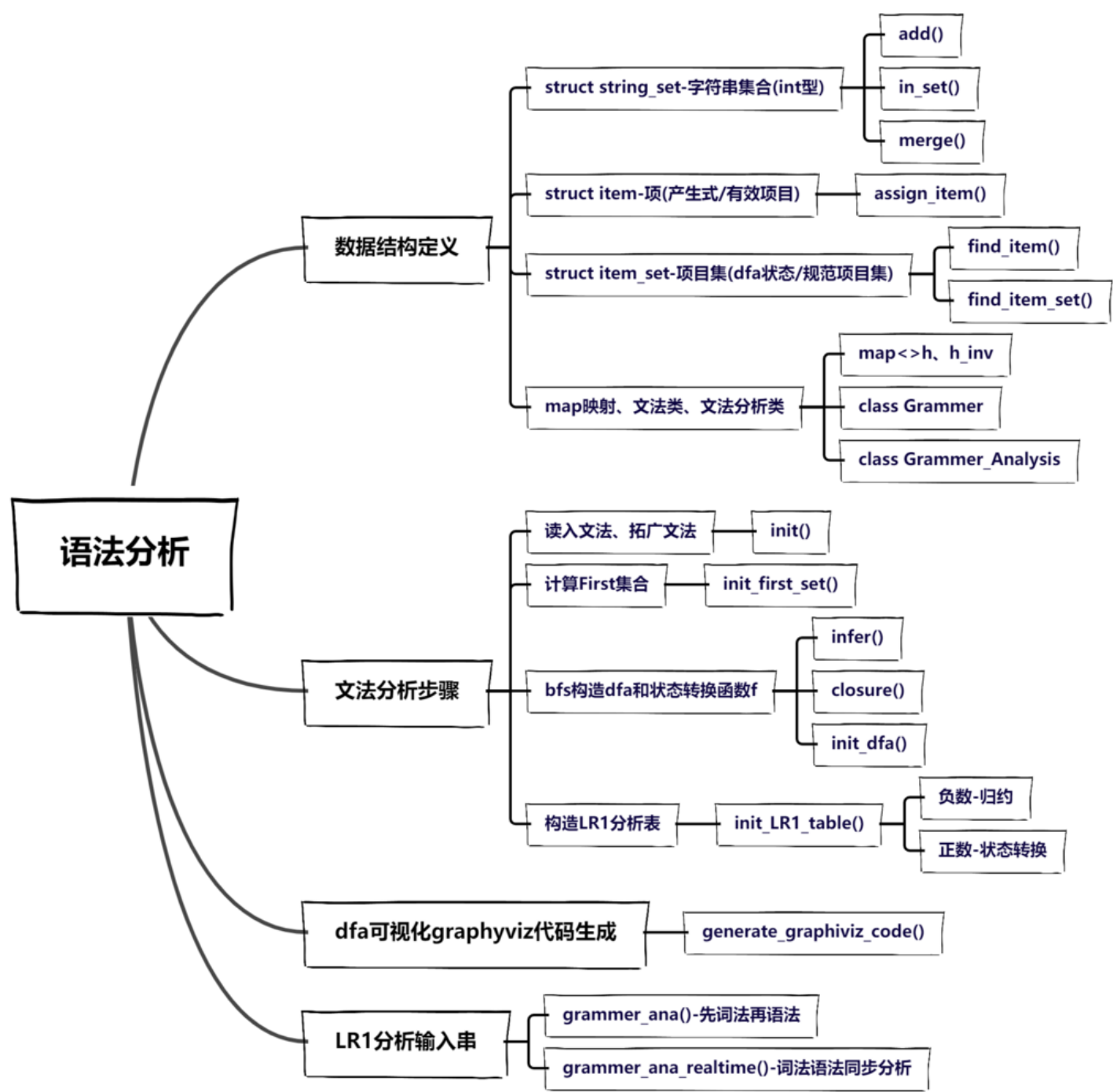
`lexical_ana_single()` 词法分析每次识别单独一个 token

`id.txt` 包括一些关键字、标识符和常数，也有标识符、常数识别的词法错误

op.txt 包括几乎所有此法定义的算术、关系和逻辑运算符

str.txt 包括单引、双引字符(串)和注释的识别，以及相应错误的识别

语法分析



语法分析思维导图

实现内容

我从输入文法（支持任意文法）开始完整构造编写 LR（1）分析程序，实现了包括但不限于以下内容：

- 1. 编程实现文法的读入，非终结符、终结符划分，产生式读入，空字符定义并构造拓广文法（保证归约时单入口）；
- 2. 编程实现了文法 First 集、Follow 集（可用于构造 LL1 select 表）的计算；
- 3. 编程实现了自底向上文法活前缀 DFA 的构造包括构造所有的规范项目集以及项目规范集间的状态转换函数 f；
- 4. 编程实现了 LR1 分析表的构造并可视化打印；
- 5. 编程实现了任意输入表达式的基于 LR1（1）分析表的分析，并且可以和词法分析程序同步进行或者先词法分析再语法分析，并可视化打印分析栈、未读字符、Action（移进/归约 包括所用产生式）等；
- 6. 编写脚本输出文法活前缀 DFA（即项目规范集合）的可视化 graphviz 代码，得到 DFA 的可视化图

总结：本次实现我基于 C++ 独立完整编程实现了 LR1 分析表从构造到使用的所有过程，并能规范地可视化打印包括 LR1 分析表、LR1 分析过程，且支持并不限于本算术表达式文法的任意自定义文法的相应构造和分析。

测试

grammer_ana () 先完整扫描输入表达式进行词法分析再进行语法分析

grammer_ana_realtime () 同步进行词法分析和语法分析

G.txt 和 G_exp.txt 实验要求的算术表达式文法和构造的待分析算术表达式字符串

G1_txt 和 G1_exp.txt 一个其他文法和对应的待分析字符串