

lab1 实验报告

学号 PB20051150 姓名 陆子睦

实验要求

本次实验需要完成一个完整的 Cminus-f 解析器。通过完善 `src/parser/lexical_analyzer.l` 和 `src/parser/syntax_analyzer.y`，从而实现基于 `flex` 的词法分析器和基于 `bison` 的语法分析器。

实验难点

感觉本次实验的第一个难点是实验中用到的 `flex` 和 `bison` 是之前没有用过的，所以需要学习如何使用这两个工具。通过自己阅读相关的文档，逐渐学会它们的使用方法。

第二个难点是理解代码框架。我们需要仔细阅读实验提供的代码框架，理解已给的一些函数以及数据结构的用法，并根据自己的理解来填写两个文件中需要补全的部分。通过对 `syntax_tree` 的理解，完成 `bison` 部分的代码。

第三个难点是词法的匹配。有很多需要注意的细节，包括如何维护 `lines`, `pos_start`, `pos_end` 等，以及需要主要 `.` 这样的符号需要在前面加上 `\` 才能匹配小数点（我就在这里卡了好久）。

第四个难点是 `bison` 部分的补全。主要是语法表达比较长，如果一点抄错了就会比较麻烦。而我在 `empty` 的处理上出现了一点问题，因为一开始没有看到 `node` 函数里面关于 `epsilon` 的处理，导致总是一个位置出现两个 `epsilon`。

第五个难点主要是我自己的问题，由于我在 Windows 下使用 `git` 克隆了仓库，导致所有的换行都是 `\r\n`，而对比的文件是 Linux 下的，所以换行是 `\n`，导致在语法树没有肉眼可见的不同的情况下每一行都不对。后来我在 docker 内部克隆了仓库，才解决了问题。

实验设计

本实验主要需要完成的就是 `src/parser/lexical_analyzer.l` 和 `src/parser/syntax_analyzer.y` 的 TO DO 部分，其实理解了还是非常直接的。

首先完成 `src/parser/lexical_analyzer.l`，会发现需要一些宏，如 `ADD` 等，而这些宏是在 `src/parser/syntax_analyzer.y` 中通过 `%token` 定义的，所以先在 `src/parser/syntax_analyzer.y` 中完成这部分的定义。通过查看 `new_syntax_tree_node()` 函数的返回值，可以看出 `union` 中 `node` 的数据类型应该是 `struct _syntax_tree_node *node;`（虽然不知道为什么不能写 `syntax_tree_node *node;`）

然后只要按照词法的规则填写就可以了，先填写一些固定的如 `else`、`+` 这样的字符串，用 `""` 包起来直接匹配就可以了。然后是用正则式定义 `ID`, `INTEGER`, `FLOAT` 这些词法。至于后面的操作，只要维护一下 `pos_start`, `pos_end`, `lines`，通过 `pass_node` 传递给 `bison`，再返回 `token` 类型就可以了。

之后是 `src/parser/syntax_analyzer.y` 的部分。可以参考 `Basics.md` 文档中的语法式子，然后添加操作，操作也就是建立节点和插入节点的孩子，只需要调用已有的函数 `node()` 就可以了。

这样如果没有什么错误的话程序就可以运行了。

实验结果验证

验证代码 1:

```
/* comment*/
/*****/
/*
*****
*/
int main(void) {
    /* comment */
/*****/
/*
*****
*/
    int array[1];
    array[1] = 0;
    return 0;
    /* comment */
/*****/
/*
*****
*/
}
```

主要用于检测注释识别是否有问题。输出如下，可以看出是正确的。

```
>--+ program
| >--+ declaration-list
| | >--+ declaration
| | | >--+ fun-declaration
| | | | >--+ type-specifier
| | | | | >--* int
| | | | | >--* main
| | | | | >--* (
| | | | | >--+ params
| | | | | | >--* void
| | | | | >--* )
| | | | | >--+ compound-stmt
| | | | | | >--* {
| | | | | | >--+ local-declarations
| | | | | | | >--+ local-declarations
| | | | | | | | >--* epsilon
| | | | | | | >--+ var-declaration
| | | | | | | | >--+ type-specifier
| | | | | | | | | >--* int
| | | | | | | | | >--* array
| | | | | | | | | >--* [
| | | | | | | | | >--* 1
| | | | | | | | | >--* ]
| | | | | | | | | >--* ;
| | | | | | >--+ statement-list
| | | | | | >--+ statement-list
| | | | | | | >--+ statement-list
| | | | | | | | >--* epsilon
```

```

| | | | | >--+ statement
| | | | | | >--+ expression-stmt
| | | | | | | >--+ expression
| | | | | | | | >--+ var
| | | | | | | | | >--+ array
| | | | | | | | | >--+ [
| | | | | | | | | >--+ expression
| | | | | | | | | | >--+ simple-expression
| | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | >--+ term
| | | | | | | | | | | >--+ factor
| | | | | | | | | | | >--+ integer
| | | | | | | | | | | >--+ 1
| | | | | | | | | | | >--+ ]
| | | | | | | | | | >--+ =
| | | | | | | | | | >--+ expression
| | | | | | | | | | | >--+ simple-expression
| | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | >--+ term
| | | | | | | | | | | >--+ factor
| | | | | | | | | | | >--+ integer
| | | | | | | | | | | >--+ 0
| | | | | | | | | | >--+ ;
| | | | | | >--+ statement
| | | | | | | >--+ return-stmt
| | | | | | | | >--+ return
| | | | | | | | >--+ expression
| | | | | | | | | >--+ simple-expression
| | | | | | | | | >--+ additive-expression
| | | | | | | | | >--+ term
| | | | | | | | | >--+ factor
| | | | | | | | | >--+ integer
| | | | | | | | | >--+ 0
| | | | | | | | >--+ ;
| | | | | | >--+ }

```

验证代码 2:

```

int main(void)
{
    int a; int b; int c;
    if (a == 1) {
        b = c;
    } else if (a > 1) {
        b = 0;
    } else
        if (a == 3)
            c = 0;
        else
            a = c;
    return 0;
}

```

主要用于验证 if else, 结果如下

```

>--+ program
| >--+ declaration-list
| | >--+ declaration
| | | >--+ fun-declaration
| | | | >--+ type-specifier
| | | | | >--+* int
| | | | | >--+* main
| | | | | >--+* (
| | | | | >--+ params
| | | | | >--+* void
| | | | | >--+* )
| | | | >--+ compound-stmt
| | | | | >--+* {
| | | | | >--+ local-declarations
| | | | | | >--+ local-declarations
| | | | | | | >--+ local-declarations
| | | | | | | | >--+ local-declarations
| | | | | | | | | >--+* epsilon
| | | | | | | | | >--+ var-declaration
| | | | | | | | | >--+ type-specifier
| | | | | | | | | | >--+* int
| | | | | | | | | | >--+* a
| | | | | | | | | | >--+* ;
| | | | | | | | >--+ var-declaration
| | | | | | | | >--+ type-specifier
| | | | | | | | | >--+* int
| | | | | | | | >--+* b
| | | | | | | | >--+* ;
| | | | | | | >--+ var-declaration
| | | | | | | >--+ type-specifier
| | | | | | | | >--+* int
| | | | | | | >--+* c
| | | | | | | >--+* ;
| | | | | >--+ statement-list
| | | | | | >--+ statement-list
| | | | | | | >--+ statement-list
| | | | | | | | >--+* epsilon
| | | | | | | >--+ statement
| | | | | | | >--+ selection-stmt
| | | | | | | | >--+* if
| | | | | | | | >--+* (
| | | | | | | | >--+ expression
| | | | | | | | | >--+ simple-expression
| | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | >--+ term
| | | | | | | | | | | >--+ factor
| | | | | | | | | | | >--+ var
| | | | | | | | | | | | >--+* a
| | | | | | | | | | | >--+ relop
| | | | | | | | | | | >--+* ==
| | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | >--+ term
| | | | | | | | | | | >--+ factor
| | | | | | | | | | | >--+ integer
| | | | | | | | | | | >--+* 1
| | | | | | | | | >--+* )
| | | | | | | >--+ statement

```

```

| | | | | | | | | | >--+ compound-stmt
| | | | | | | | | | | >--* {
| | | | | | | | | | | >--+ local-declarations
| | | | | | | | | | | | >--* epsilon
| | | | | | | | | | | >--+ statement-list
| | | | | | | | | | | | >--+ statement-list
| | | | | | | | | | | | | >--* epsilon
| | | | | | | | | | | | >--+ statement
| | | | | | | | | | | | | >--+ expression-stmt
| | | | | | | | | | | | | | >--+ expression
| | | | | | | | | | | | | | | >--+ var
| | | | | | | | | | | | | | | | >--* b
| | | | | | | | | | | | | | | | >--* =
| | | | | | | | | | | | | | | | >--+ expression
| | | | | | | | | | | | | | | | >--+ simple-expression
| | | | | | | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | | | | | | | >--+ term
| | | | | | | | | | | | | | | | | >--+ factor
| | | | | | | | | | | | | | | | | >--+ var
| | | | | | | | | | | | | | | | | >--* c
| | | | | | | | | | | | | | | | >--* ;
| | | | | | | | | | | | | | | | >--* }
| | | | | | | | | | | >--* else
| | | | | | | | | | | >--+ statement
| | | | | | | | | | | | >--+ selection-stmt
| | | | | | | | | | | | | >--* if
| | | | | | | | | | | | | >--* (
| | | | | | | | | | | | | >--+ expression
| | | | | | | | | | | | | | >--+ simple-expression
| | | | | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | | | | | >--+ term
| | | | | | | | | | | | | | | >--+ factor
| | | | | | | | | | | | | | | >--+ var
| | | | | | | | | | | | | | | | >--* a
| | | | | | | | | | | | | | | >--+ relop
| | | | | | | | | | | | | | | | >--* >
| | | | | | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | | | | | >--+ term
| | | | | | | | | | | | | | | >--+ factor
| | | | | | | | | | | | | | | >--+ integer
| | | | | | | | | | | | | | | >--* 1
| | | | | | | | | | | | | >--* )
| | | | | | | | | | | | >--+ statement
| | | | | | | | | | | | | >--+ compound-stmt
| | | | | | | | | | | | | | >--* {
| | | | | | | | | | | | | | >--+ local-declarations
| | | | | | | | | | | | | | | >--* epsilon
| | | | | | | | | | | | | | >--+ statement-list
| | | | | | | | | | | | | | >--+ statement-list
| | | | | | | | | | | | | | | >--* epsilon
| | | | | | | | | | | | | | >--+ statement
| | | | | | | | | | | | | | >--+ expression-stmt
| | | | | | | | | | | | | | | >--+ expression
| | | | | | | | | | | | | | | >--+ var
| | | | | | | | | | | | | | | | >--* b
| | | | | | | | | | | | | | | | >--* =
| | | | | | | | | | | | | | | >--+ expression
| | | | | | | | | | | | | | | >--+ simple-expression

```

																>--+ additive-
expression																>--+ term
																>--+ factor
																>--+ integer
																>--* 0
																>--* ;
																>--* }
																>--* else
																>--+ statement
																>--+ selection-stmt
																>--* if
																>--* (
																>--+ expression
																>--+ simple-expression
																>--+ additive-expression
																>--+ term
																>--+ factor
																>--+ var
																>--* a
																>--+ relop
																>--* ==
																>--+ additive-expression
																>--+ term
																>--+ factor
																>--+ integer
																>--* 3
																>--*)
																>--+ statement
																>--+ expression-stmt
																>--+ expression
																>--+ var
																>--* c
																>--* =
																>--+ expression
																>--+ simple-expression
																>--+ additive-expression
																>--+ term
																>--+ factor
																>--+ integer
																>--* 0
																>--* ;
																>--* else
																>--+ statement
																>--+ expression-stmt

```

| | | | | | | >--+ return-stmt
| | | | | | | | >--* return
| | | | | | | | >--+ expression
| | | | | | | | | >--+ simple-expression
| | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | >--+ term
| | | | | | | | | | | | >--+ factor
| | | | | | | | | | | | >--+ integer
| | | | | | | | | | | | >--* 0
| | | | | | | | >--* ;
| | | | | | >--* }

```

验证代码 3:

```

void Fibonacci(int n){
    int a=0; int b=1; int c;
    if(n>0){
        c = a + b;
        a = b;
        b = c;
        Fibonacci(n-1);
    }
}

int main(){
    int n;
    Fibonacci(n)
    return 0;
}

```

Fibonacci 数列代码，结果如下:

```

>--+ program
| >--+ declaration-list
| | >--+ declaration
| | | >--+ fun-declaration
| | | | >--+ type-specifier
| | | | | >--* int
| | | | | >--* main
| | | | | >--* (
| | | | | >--+ params
| | | | | | >--* void
| | | | | >--* )
| | | | | >--+ compound-stmt
| | | | | | >--* {
| | | | | | >--+ local-declarations
| | | | | | | >--+ local-declarations
| | | | | | | | >--+ local-declarations
| | | | | | | | | >--* epsilon
| | | | | | | | | >--+ var-declaration
| | | | | | | | | | >--+ type-specifier
| | | | | | | | | | | >--* int
| | | | | | | | | | | >--* a
| | | | | | | | | | | >--* ;
| | | | | | | | >--+ var-declaration

```

```

| | | | | | | >--+ type-specifier
| | | | | | | | >--* int
| | | | | | | | >--* b
| | | | | | | | >--* ;
| | | | | | | >--+ var-declaration
| | | | | | | | >--+ type-specifier
| | | | | | | | | >--* int
| | | | | | | | >--* c
| | | | | | | | >--* ;
| | | | | | >--+ statement-list
| | | | | | | >--+ statement-list
| | | | | | | | >--+ statement-list
| | | | | | | | | >--* epsilon
| | | | | | | | >--+ statement
| | | | | | | | | >--+ selection-stmt
| | | | | | | | | | >--* if
| | | | | | | | | | >--* (
| | | | | | | | | | >--+ expression
| | | | | | | | | | | >--+ simple-expression
| | | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | | | >--+ term
| | | | | | | | | | | | | >--+ factor
| | | | | | | | | | | | | >--+ var
| | | | | | | | | | | | | | >--* a
| | | | | | | | | | | | | >--+ relop
| | | | | | | | | | | | | | >--* ==
| | | | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | | | | >--+ term
| | | | | | | | | | | | | | >--+ factor
| | | | | | | | | | | | | | >--+ integer
| | | | | | | | | | | | | | >--* 1
| | | | | | | | | | | | | >--* )
| | | | | | | | | | | | >--+ statement
| | | | | | | | | | | | | >--+ compound-stmt
| | | | | | | | | | | | | | >--* {
| | | | | | | | | | | | | | >--+ local-declarations
| | | | | | | | | | | | | | | >--* epsilon
| | | | | | | | | | | | | | >--+ statement-list
| | | | | | | | | | | | | | | >--+ statement-list
| | | | | | | | | | | | | | | | >--* epsilon
| | | | | | | | | | | | | | | | >--+ statement
| | | | | | | | | | | | | | | | >--+ expression-stmt
| | | | | | | | | | | | | | | | | >--+ expression
| | | | | | | | | | | | | | | | | >--+ var
| | | | | | | | | | | | | | | | | | >--* b
| | | | | | | | | | | | | | | | | >--* =
| | | | | | | | | | | | | | | | | >--+ expression
| | | | | | | | | | | | | | | | | | >--+ simple-expression
| | | | | | | | | | | | | | | | | | >--+ additive-expression
| | | | | | | | | | | | | | | | | | | >--+ term
| | | | | | | | | | | | | | | | | | | >--+ factor
| | | | | | | | | | | | | | | | | | | >--+ var
| | | | | | | | | | | | | | | | | | | | >--* c
| | | | | | | | | | | | | | | | | >--* ;
| | | | | | | | | | | | | | | | | >--* }
| | | | | | | | | | | | | | >--* else
| | | | | | | | | | | | | | >--+ statement
| | | | | | | | | | | | | | >--+ selection-stmt

```


[illegible]

```

| | | | | | | | | | >--+ integer
| | | | | | | | | | >--* 3
| | | | | | | | | | >--* )
| | | | | | | | | | >--+ statement
| | | | | | | | | | >--+ expression-stmt
| | | | | | | | | | >--+ expression
| | | | | | | | | | >--+ var
| | | | | | | | | | >--* c
| | | | | | | | | | >--* =
| | | | | | | | | | >--+ expression
| | | | | | | | | | >--+ simple-expression
| | | | | | | | | | >--+ additive-expression
| | | | | | | | | | >--+ term
| | | | | | | | | | >--+ factor
| | | | | | | | | | >--+ integer
| | | | | | | | | | >--* 0
| | | | | | | | | | >--* ;
| | | | | | | | | | >--* else
| | | | | | | | | | >--+ statement
| | | | | | | | | | >--+ expression-stmt
| | | | | | | | | | >--+ expression
| | | | | | | | | | >--+ var
| | | | | | | | | | >--* a
| | | | | | | | | | >--* =
| | | | | | | | | | >--+ expression
| | | | | | | | | | >--+ simple-expression
| | | | | | | | | | >--+ additive-expression
| | | | | | | | | | >--+ term
| | | | | | | | | | >--+ factor
| | | | | | | | | | >--+ var
| | | | | | | | | | >--* c
| | | | | | | | | | >--* ;
| | | | | >--+ statement
| | | | | >--+ return-stmt
| | | | | >--* return
| | | | | >--+ expression
| | | | | >--+ simple-expression
| | | | | >--+ additive-expression
| | | | | >--+ term
| | | | | >--+ factor
| | | | | >--+ integer
| | | | | >--* 0
| | | | | >--* ;
| | | | | >--* }

```

实验反馈

本次实验中，我学会了 `flex` 和 `bison` 的使用方法，同时又锻炼了排除 `bug` 的能力，感觉得到了很多收获。