

- 学号: 221240093
- 姓名: 陈力峥
- 邮箱: 221240093@smail.nju.edu.cn

文件结构

```
/
/Code
    Makefile
    lexical.l
    syntax.y
    main.c

/Test
    test1.cmm
    test2.cmm

/report.pdf
/parser
/README
```

编译

编译实验一采用 `Makefile`，在项目根目录下执行 `make` 即可，随后生成 `parser` 可执行文件。

实验内容

完成实验一必做以及选做部分：

- C++ 语言的词法分析，包含 `Appendix.pdf` 中给出的 C++ 语言的所有词法单元，特别的，支持八进制、十六进制数字、指数型浮点数以及嵌套注释的识别。
- C++ 语言的语法分析，包含 `Appendix.pdf` 中给出的 C++ 语言的所有语法单元，支持部分错误恢复机制。

实验细节

- 对于基本的词法单元以及高级的词法单元，采用 `Node` 结构体进行存储(即都声明为 `Node` 类型)，其格式如下：

```
typedef struct Node {  
    char *name;  
    int line;  
    NodeValue value;  
    struct Node *next;  
    struct Node *child;  
} Node;
```

各个字段分别存储了该词法单元的名称、行号、值以及指向下一个词法单元的指针以及指向子节点的指针。其中 `NodeValue` 类型用于 `int`, `float`, `string` 类型的值的存储，而其他类型的 `node` 则不设置 `value` 的值。

- 对于嵌套注释的识别，采用 `state` 的方式进行处理，当识别到 `/*` 时，进入 `IN_COMMENT` 状态，随后只匹配当前状态对应的匹配规则，当识别到 `*/` 时，退出 `COMMENT` 状态，此时不会将 `COMMENT` 词法单元加入到词法单元链表中。
- 语法规则在 `Appendix.pdf` 的基础上，增加各操作符的结合性，同时通过定义 `LOWER_THAN_ELSE` 和 `ELSE` 的优先级，解决了 `if-else` 语句的移入/归约二义性问题。