

The compiler converts a high-level language to a low-level language.
After the high-level language goes through lexical analysis we get tokens.

HLL \rightarrow Lexical analysis \longrightarrow Tokens

int x;
INT ID SEMICOLON

For all the lexemes of the high-level language, we get tokens after lexical analysis.

After the tokens go through syntax analysis we get a parse tree.

According to the given grammar, the rules will be checked. If the rules get matched we generate the parse tree and say the tokens to be synthetically organised. In short, if the tokens are synthetically organised finally a parse tree gets generated.

Tokens \rightarrow Syntax analysis \longrightarrow Parse Tree

For this lab, if the nodes are grammatically correct then we will print them.

The grammar contains the necessary rules that a function needs to run.

$$y = x + x$$

The grammar given is for a function. But this is not a function, also y and x is not defined and the semi colon is missing. So this is grammatically incorrect.

At line 1 syntax error

So this will be printed.

```

C symbol_info.h      C TreeNode.h      syntax_analyzer.y
C: > Users > Asus > Desktop > Lab 03 Final > syntax_analyzer.y
10
11     extern int yyparse(void);
12     extern int yylex(void);
13
14     extern FILE *yyin;
15
16     extern YYSTYPE yyval;
17
18     std::ofstream outlog;
19
20     int lines = 1;
21
22     TreeNode* rootNode = nullptr;
23
24     void yyerror(const char *s)
25     {
26         outlog << "At line " << lines << " " << s << std::endl << std::endl;
27     }
28
29 %}
30
31 %token IF ELSE FOR WHILE DO BREAK INT CHAR FLOAT DOUBLE VOID RETURN SWITCH CASE DEFAULT CONTINUE PRINTF ADDOP MUL
32

```

A root node variable is created where the reference of the root node is saved. (root node = first node) so that the tree can be traversed in post order.

```

C symbol_info.h      C TreeNode.h      syntax_analyzer.y
C: > Users > Asus > Desktop > Lab 03 Final > C TreeNode.h
1 #ifndef TREE_NODE_H
2 #define TREE_NODE_H
3
4 #include <string>
5 #include <vector>
6
7 class TreeNode {
8 private:
9     std::string type;
10    std::string value;
11    std::vector<TreeNode*> children;
12
13 public:
14     TreeNode(const std::string& type, const std::string& value = "")
15         : type(type), value(value) {}
16
17     static TreeNode* createNonTerminalNode(const std::string& type) {
18         return new TreeNode(type);
19     }
20
21     static TreeNode* createTerminalNode(const std::string& type, const std::string& value = "") {
22         return new TreeNode(type, value);
23     }

```

In type, the tokens are stored

Value = lexeme → things that are read from the input file.

```

C symbol_info.h      C TreeNode.h      syntax_analyzer.y
C: > Users > Asus > Desktop > Lab 03 Final > syntax_analyzer.y
32
33     %nonassoc LOWER_THAN_ELSE
34     %nonassoc ELSE
35
36     %start program
37
38     %%
39
40     program : program unit
41     {
42         TreeNode* node = TreeNode::createNonTerminalNode("program");
43         node->addChild($1->getNode());
44         node->addChild($2->getNode());
45         $$ = new symbol_info("program", "non_terminal", node);
46         rootNode = node;
47     }
48     | unit
49     {
50
51     }
52     ;
53
54     unit : var_declaration

```

As `createNonTerminalNode` is called, inside the `TreeNode.h` file, a non-terminal node will be created.

When `addChild` method is called, it adds the child under the respective node on which it was called upon.

```

64     func_definition : type_specifier id_name LPAREN parameter_list RPAREN compound_statement
65     {
66         TreeNode* node = TreeNode::createNonTerminalNode("func_definition"); // WILL THERE
67         node->addChild($1->getNode());
68         node->addChild($2->getNode());
69         node->addChild(TreeNode::createTerminalNode("LPAREN", "("));
70         node->addChild($4->getNode());
71         node->addChild(TreeNode::createTerminalNode("RPAREN", ")"));
72         node->addChild($6->getNode());
73         $$ = new symbol_info("func_definition", "non_terminal", node); // WILL THERE
74     }

```

Node contains the reference of the tree node class object

```

C symbol_info.h      C TreeNode.h      syntax_analyzer.y
C: > Users > Asus > Desktop > Lab 03 Final > C TreeNode.h
14     TreeNode(const std::string& type, const std::string& value = "") :
15         type(type), value(value) {}
16
17     static TreeNode* createNonTerminalNode(const std::string& type) {
18         return new TreeNode(type);
19     }
20
21     static TreeNode* createTerminalNode(const std::string& type, const std::string& value = "") {
22         return new TreeNode(type, value);
23     }
24
25     void addChild(TreeNode* child) {
26         children.push_back(child);
27     }
28
29     const std::string& getType() const {
30         return type;
31     }
32
33     const std::string& getValue() const {
34         return value;
35     }

```

Getnode method returns the already created branch which basically gets added as a child.

$\$ \$ \leftarrow$ symbol table

```
36
40     program : program unit
41     {
42         TreeNode* node = TreeNode::createNonTerminalNode("program");
43         node->addChild($1->getNode());
44         node->addChild($2->getNode());
45         $ = new symbol_info("program", "non_terminal", node);
46         rootNode = node;
```

Symbol info class is created and is added as a pointer in the symbol table($\$ \$$)

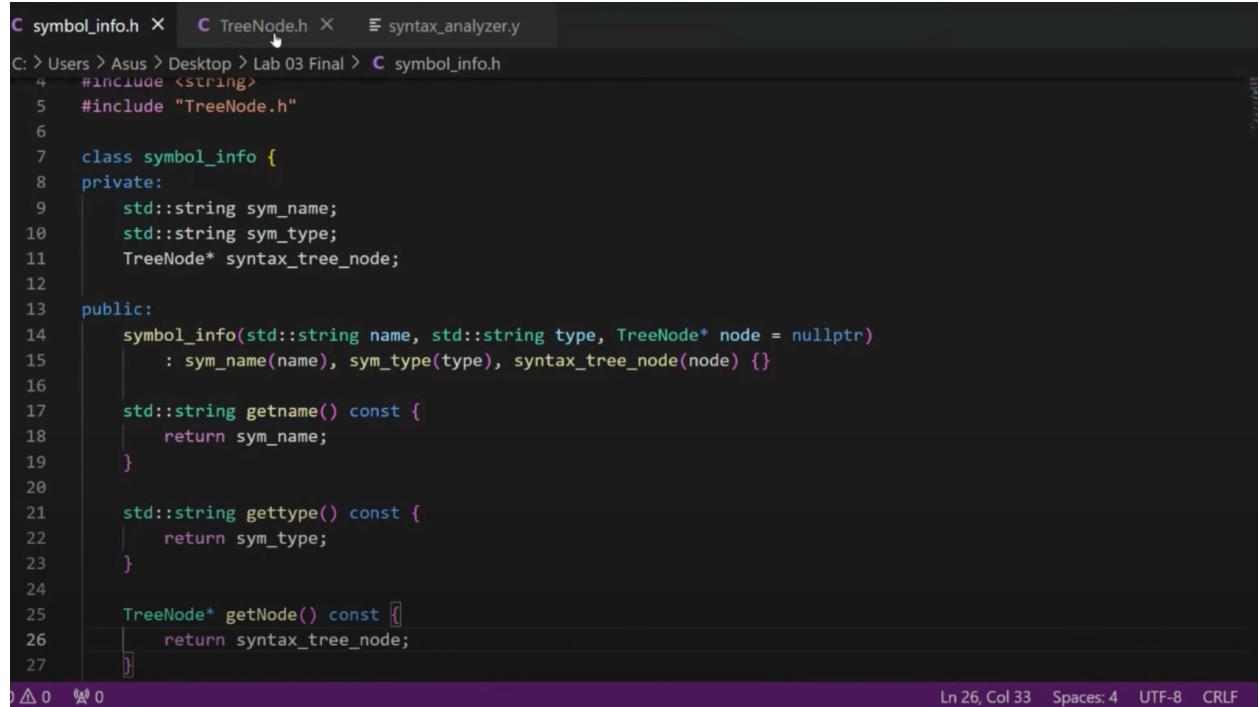
```
369     if (rootNode) {
370         rootNode->postOrderTraversal(outlog);
371     }
```

Outlog parameter is passed which is the output file to write the value i.e. the lexeme that was read in the input file.

```
50
51     if (!value.empty()) {
52         outFile << " " << value;
53     }
```

If the node is not empty then values will be printed along with space.

To summarize we are creating nodes based on terminal and non-terminal. After that, we add children by calling addChild method. Then getNode method is used to add the already created branch. Finally, the symbol info class is created and is added under the symbol info table.



The screenshot shows a code editor window with three tabs at the top: 'symbol_info.h' (active), 'TreeNode.h', and 'syntax_analyzer.y'. The current file, 'symbol_info.h', contains the following C++ code:

```
C: > Users > Asus > Desktop > Lab 03 Final > C symbol_info.h
4  #include <string>
5  #include "TreeNode.h"
6
7  class symbol_info {
8  private:
9      std::string sym_name;
10     std::string sym_type;
11     TreeNode* syntax_tree_node;
12
13 public:
14     symbol_info(std::string name, std::string type, TreeNode* node = nullptr)
15         : sym_name(name), sym_type(type), syntax_tree_node(node) {}
16
17     std::string getname() const {
18         return sym_name;
19     }
20
21     std::string gettype() const {
22         return sym_type;
23     }
24
25     TreeNode* getNode() const {
26         return syntax_tree_node;
27     }
}
```

At the bottom of the editor, there are status indicators: 'Δ 0', '0 0', 'Ln 26, Col 33', 'Spaces: 4', 'UTF-8', and 'CRLF'.

Syntax tree node contains the reference of the symbol info class.