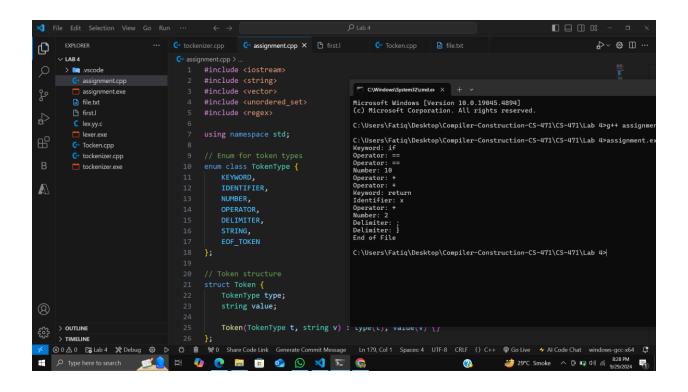
Registration No: 2021-CS-140

Instructor: Mr. Laeeq Khan Niazi

CS-471 Lab 4

Compiler Construction

By using best optimized data structures write the tokenizer for your own language.



```
	imes File Edit Selection View Go Run \cdots \leftarrow 	o
                                                                                                         $>∨ ∰ Ⅲ ...
                       file.txt
                            C+ assignment.cpp > ☐ TokenType > ☐ EOF_TOKEN
     > 💌 .vscode
                             29 class Tokenizer {
                                    vector<Token> tokenize(const string& code) {
       assignment.exe
       file.txt
       first.l
$
                                            if (isDelimiter(code[pos])) {
       lexer.exe
                                                tokens.emplace_back(TokenType::DELIMITER, string(1, code[pos]));
       C Tocken.cpp
       tockenizer.cpp
if (code[pos] == '"') {
                                               smatch match;
                                                string remainingCode = code.substr(pos);
                                                if (regex_search(remainingCode, match, stringRegex)) {
                                                   tokens.emplace_back(TokenType::STRING, match.str());
                                                   pos += match.length();
                                            cout << "Unexpected character: " << code[pos] << endl;</pre>
> OUTLINE
                                            pos++;
    > TIMELINE

    39°C Smoke
    △ ⊕ 📭 Φ)

    40°C Smoke
    △ ⊕ 📭 Φ)

ff 🔑 Type here to search 📈 🏮 🖟 🕡 📵 📋 💇 🔘 刘 🖫 🦙
                                                                                      (3)

★ File Edit Selection View Go Run …

                       ... C+ tockenizer.cpp C+ assignment.cpp X C first.l
    EXPLORER
                                                                    ← Tocken.cpp  file.txt
                                                                                                                  ♪∨ ⊜ □ …
0
    V LAB 4
                            👉 assignment.cpp > 🚭 TokenType > 🚭 EOF_TOKEN
     > xscode
                                   void printTokens() {
وع
       assignment.exe
                                                   break:
       file.txt
       first.l
                                                case TokenType::EOF_TOKEN:
    cout << "End of File" << endl;</pre>
string code = R"(if (x == 10) { return x + 2; })";
                                     Tokenizer tokenizer;
                                     vector<Token> tokens = tokenizer.tokenize(code);
                                     // Print the tokens
                                     tokenizer.printTokens();
> outline
                                     return 0;
    > TIMELINE
   ⊗ 0 🛆 0 🔁 lab 4 🛠 Debug 🕲 ▷ 🌣 🔞 💖 0 Share Code Link Generate Commit Message Ln 17, Col 14 Spaces: 4 UTF-8 CRLF () C++ 🖗 Go Live 🗲 Al Code Chat windows-gcc-x64 🕻
   ## 29°C Smoke ^ @ ₽ Ф1) // 829 PM
                                                                                      ?
```