COMSATS Institute of
Information Technology

Group Members:

M. Abdullah Arshad (SP20-BCS-033)

Hasnain Ahmed (FA20-BCS-005)

Class/Section: BCS-7 (A)

Subject: CC-Lab (Compiler Construction)

Submission To: Sir Bilal Haider

Date: 28-Dec-2023

CC-Lab Terminal:

Question 1: Write brief of the project.

Answer: We have develop a mini compiler that performs Lexical Analysis and LR parsing of strings.

- **Lexical analyzer:** It breaks down the source code into a sequence of tokens. Tokens are the smallest units of meaning in a programming language, such as keywords, identifiers, literals, and operators.

  The main tasks of a lexical analyzer include:

- **Tokenization:** Breaking the source code into tokens based on the language's syntax rules.
- **Removing Whitespace and Comments:** Discarding elements like spaces, tabs, and comments that do not contribute to the meaning of the program.

  The specific actions performed by a lexical analyzer can vary depending on the programming language. Here's a general overview of what a lexical analyzer does in C#:
  - **Scanning:** Reads the source code character by character.
  - **Lexical Error Detection:** Identifies and reports lexical errors, such as misspelled keywords or undefined symbols.
  - **Token Generation:** Recognizes and generates tokens for keywords, identifiers, literals, operators, and other language constructs.

❖ **LR Parsing:** It based on a shift-reduce approach, where the parser shifts input symbols onto a stack until it identifies a sequence that can be reduced to a grammar production.

LR parsers main tasks include:

- **Shift Operation**: The parser shifts (moves) input symbols onto a stack until it identifies a valid right-hand side of a grammar production.
- **Reduce Operation**: Once a valid right-hand side is on top of the stack, the parser replaces that sequence with the corresponding non-terminal symbol of the grammar production. This is known as a reduce operation.
- **Acceptance**: The process continues until the parser accepts the entire input, indicating that the input adheres to the grammar rules.

-End