**University of Central Punjab**

**Faculty of Information Technology**



**Compiler Construction**

Project (Phase-1) (Section All) Spring (2025)

Submission Before: 11:00PM - 6-29-2025

UCP-Compiler:

In lexical following token types required to be implemented.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Detail | Example |
| 1 | Identifiers | Start from \_ or Letter and contains at least one “\_”. Other rules for ID will remain same. | \_rate2, \_rate\_, rat1e2 …etc |
| 2 | Numbers | [+-]?(D+)(\.D+)?, and exponent numbers. | 3.43433E+13, |
| 3 | Operators | !=, <>, =:= , ==, \*, +, /,-, >>,<<, ++, =+, &&, ||, =>, =<, %, :, :: , -- |  |
| 4 | Punctuations | [,{,< , >,},] |  |
| 5 | Keyword | loop, agar, magar, asm, else, new, this, auto, enum, operator, throw, bool , explicit, private, true, break, export, protected , try, case, extern, public, typedef, catch, false, register, typeid, char, float, typename,  class, for, return, union, const, friend, short, unsigned  goto, signed, using, continue, if, sizeof, virtual, default, inline, static, void, delete, int, volatile, do, long, struct,  double, mutable, switch, while, namespace |  |

# Assignment Description:

For this assignment, you have to implement a **lexical analyzer**, also called a scanner. This assignment includes following parts:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **PARTS** | **Output** | **Marks** | **Deadline** |
| **1** | Generate FA for each regular expression | Document | 25 | 6 Dec |
| **2** | Code - Table Driven Approach | Source Code Files | 100 | 6 Dec |

# Note: (Late will be penalty of deduction of 25% of total marks per day)

# Tools:

Language (For Development): C++/Python

**Note: Student cannot use built-in data structure. Student can use his own data structure Hash Table, Linked List which he/she developed in data structure course. In this case student has to show me the code before using it.**

# Evaluating Criteria:

1. Source code should reflect the detail given in documents (other parts).
2. A text file with valid source code will be input of the scanner and Token file will be output of the scanner tool.
3. All points discussed in the class related to scanner implementation should reflect in the assignment.
4. Application should not do which is not supposed to do.

**SCANNER**

Token.txt

Sourcecode.txt

Error.txt

1. In case of invalid source code tool must generate error list.

**GOOD LUCK**