**Compiler Construction**

**Assignment 4a Fall 2020 Net Weightage: TBD**

**Due: Monday, November 10, 2022 11:59 pm (Non-extendable)**

* **Late Submission will not be accepted for any reason.**
* There are no marks for empty submission.
* **You have to implement this assignment in a programming language chosen for assignment 1**
* Please submit a .CPP/.PY file. The name of the file will be RollNumber\_3.cpp/py.
* Avoid all sorts of plagiarism as it can have serious consequences.
* In case of any queries, feel free to email me at [bscs17046@itu.edu.pk](mailto:bscs17046@itu.edu.pk)

**Instructions**

You are required to write a C++/Python program to calculate first and follow sets of a CFG, and eventually, compute a top-down LL(1) parsing table.

1. Your code must read the grammar from a text file. The format of the grammar file is as follows;

* Every production is written in a single line, where each grammar symbol is separated with a tab.
* The first symbol of each line is a non-terminal representing the left hand side of the production whereas the rest of the live represent the right hand side of the production.
* the tilde character (~) represents epsilon i.e. the empty character.

1. Your grammar does not have any direct or indirect left factoring or left recursion.
2. The right-hand side is restricted to at max 10 literals.
3. The non-terminals are composed of capital letters only, whereas the terminals can be any combination of lower or uppercase letters, digits, operators, or the underscore character.
4. Only those symbols that appear on the left-hand side of any of the productions are the grammar symbol (i.e. Non-Terminals) whereas the rest of the symbols are the terminals.
5. Non-Terminal appeared on the left-hand side of the first production is the start symbol.

**Problem Statement:**

**Compute First Set**

You are required to write a program that accepts a grammr.txt file as defined above and compute First sets of all the non-terminals and write them into a text file called First.txt. The format of the First.txt is such that each line starts with the non-terminal and its first set appears in the same line such that each element is tab separated.

**Compute Follow Set**

You are required to write a program that accepts a grammr.txt and the First.txt file as defined above and compute Follow sets of all the non-terminals and write them into a text file called Follow.txt. The format of the Follow.txt is such that each line starts with the non-terminal and its follow set appears in the same line such that each element is tab separated. Use the dollar character ($) as the sentinel terminal character.

**Compute the LL(1) Parsing table**

Your program must accept the grammar.txt, First.txt and Follow.txt files and compute the LL(1) parsing table as discussed in the lectures and as written in your text book. Rather than writing the full grammar production you can write the production number into your parsing table entries. The parsing table must be written in an excel file called LL1ParsingTable.xls such that first column holds the non-terminals, first row holds the terminals (the first cell remains empty for table reference). Each element of the parsing table is empty, or a number representing a production. Consult your text book page 225 for the parsing table formatting reference.

**Note:** The sample input and output will be provided to you by the course TA.