National University of Computer & Emerging Sciences, Karachi

Spring-2018 CS-Department

MidTerm 1

23rd February 2018, 10:30 am – 11:30am

Course Code:	**Course Name:** Compiler Construction
Instructor Name / Names: M. Shahzad	
Student Roll No:	**Section No:**

Instructions:

- Return the question paper.
- Read each question completely before answering it. There are **3 questions and 2 pages**.
- In case of any ambiguity, you may make assumption. But your assumption should not contradict any statement in the question paper.
- All the answers must be solved according to the sequence given in the question paper.
- This paper is subjective. All the questions should be attempted on the answer sheet.
- All questions carry equal marks and equally distributed in sub parts.

SOLUTION PAPER

Time: 60 minutes. **Max Marks**: 60 points

Question 1 (20 points = 1 * 20):

a) Select the best answer and write either A, B, C or D from the options given below in each statement:

1. What is the primary role of the parser?
   A. Verifying that the tokens can be generated by the grammar of the source language
   B. Output a set of tokens and production rules
   C. Construct the parse tree
   D. All of the above
   Correct answer: (a,c)

2. Both top-down and bottom-up parsers
   A. Can parse specific sub-classes of a given grammar
   B. Are expressive enough to describe most programming constructs
   C. Detect errors but do not support efficient detection of these errors
   D. All of the above
   Correct answer: (a,b)

3. Which of the following statements are true?
   A. Every regular language is a context-free language
   B. Every construct that can be described by a CFG can also be described by a regex
   C. CFGs are a more powerful notation than regexes
   D. All of the above
   Correct answer: (a,c)

4. Which of the following statements are true concerning ambiguity?
   A. It is possible to automate the process of converting an ambiguous grammar into an unambiguous one
   B. A set of standardized techniques is available for handling every ambiguous situation

C. Disambiguation rules can resolve the problem of every ambiguous grammar
D. None of the above
Correct answer: (d)

5. In parsing, why we might need to use the declarations %left and %right:
   A. To resolve the ambiguity related to left-most and right-most derivations respectively
   B. To resolve the ambiguity related to the associativity of some numerical operations
   C. To resolve the ambiguity related to the precedence of some numerical operations
   D. All of the above
   Correct answer: (b,c)

b) Provide 2-3 sentence replies to all of the following short questions. I will not consider the answer that exceeds 3 lines. [05 points]

1. What is the difference between a token and a pattern?

   Token is a sequence of characters that can be treated as a single logical entity. Typical tokens are, 1) Identifiers 2) keywords 3) operators 4) special symbols 5)constants.

   Pattern is a set of strings in the input that is described by a rule.

2. How are the terms "alphabet" and "regular expressions" related to each other?

   Alphabets are the symbols that is used to defined the string in a language. Whereas the regular expression defines the pattern of alphabets used in string in a language.

3. What does the term “ambiguous CFG” mean? Describe with the help of an example

   CFG is said to be ambiguous for which there exists a string that can have more than one leftmost derivation or parse tree

4. Differentiate between deterministic and non-deterministic grammar with one example.

   A **CFG** is non-deterministic (**unambiguous)** iff *every* string has *at most* one valid derivation according to the CFG.

   A CFG is non deterministic (ambiguous) iff there are at least two productions with same terminal prefix on the right side of them.

5. Explain the need of symbol table manager.

   The symbol table carries the collected information about each named object in the program to other phases of the compiler

Question 2 (20 points = 10 * 2)

Consider the following grammar:

$S \rightarrow aS \mid Ab$
$A \rightarrow XYZ \mid \varepsilon$
$X \rightarrow cS \mid \varepsilon$
$Y \rightarrow dS \mid \varepsilon$
$Z \rightarrow eS$

a) Give a leftmost derivation of the string ***aebb***.

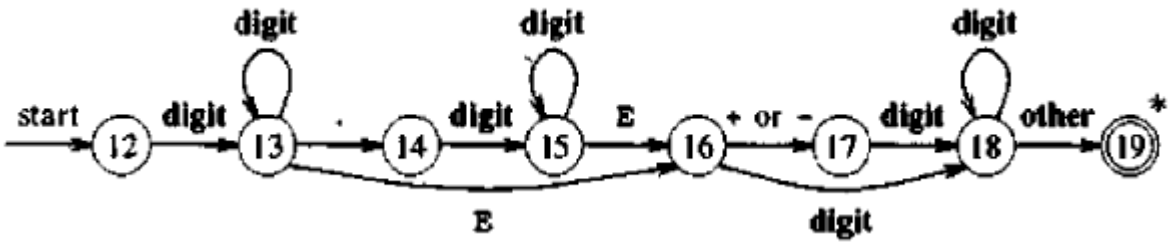
S
aS
aAb
aXYZb
aYZb
aZb
aeSb
aeAbb
aebb

b) Show the parse tree of leftmost derivation for the string given in part (a).

Question 3 (20 points = 10 * 2):

a) For the following **regex**, draw a (totally defined) DFA:

num → digit$^+$ (. digit$^+$)? (E (+|-)? digit$^+$)?



b) Find the regular expression to describe the cell numbers in Pakistan for all the various forms, also consider the optional international dialing code of Pakistan and Karachi. While giving your regular expression, you must consider the length of international dialing code(Pakistan & Karachi), mobile network code and individual cell number (7 digits). Following are the two valid sample numbers with and without international dialing code:

Sample 1: **009 2213312772751**

Sample 2: **03012526885**

BEST OF LUCK!