UNIVERSITY OF AGRICULTURE
COLLEGE OF NATURAL SCIENCES
2009/2010 FIRST SEMESTER EXAMINATION

Course Title: COMPILER CONSTRUCTION Course Code : CSC415
Instructions: Answer any **FOUR** questions. Time Allowed: 2½hrs

1a. What is a compiler?
b. What is a nondeterministic finite automaton?
c. What are the reasons for keeping the phases of compiler separate ?
d. Given the regular expression a*(a|b)aa: Construct an equivalent NFA .

2. Construct DFAs for each of the following regular languages. In all cases the alphabet is {a, b}.
a) The set of strings that has exactly 3 b's (and any number of a's).
b) The set of strings where the number of b's is a multiple of 3 (and there can be any number of a's).
c) The set of strings where the difference between the number of a's and the number of b's is a multiple of 3.
d) Construct a DFA that recognises balanced sequences of parenthesis with a maximal nesting depth of 3, e.g., e, ()(), (()(())) or (()())()() but not (((()))) or (()(()(()))).

3. Given that binary number strings are read with the most significant bit first and may have leading zeroes, construct DFAs for each of the following languages:
a) Binary number strings that represent numbers that are multiples of 4, e.g., 0, 100 and 10100.
b) Binary number strings that represent numbers that are multiples of 5, e.g., 0, 101, 10100 and 11001.
Hint: Make a state for each possible remainder after division by 5 and then add a state to avoid accepting the empty string.
c) Given a number n, what is the minimal number of states needed in a DFA that recognises binary numbers that are multiples of n? Hint: write n as $a*2^b$, where a is odd.

4. Show that regular languages are closed under prefix, suffix, subsequence and Reversal. Hint: show how an NFA N for a regular language L can be transformed to an NFA Np for the set of prefixes of strings from L, and similarly for the other operations.

5a. Draw the syntax tree for the string aabbbcc using the grammar below:

T --→ R
T --→ aTc
R --→ RbR

b. Give the derivation of the string aabbbcc using the grammar in (5a)

c. Show that the grammar

A ---→ −A
A ---→ A−id
A ---→ id

is ambiguous by finding a string that has two different syntax trees.

6a. Pick some programming language that you know well and determine which of the following objects share name spaces: Variables, functions, procedures and types. If there are more kinds of named objects (labels, data constructors, modules, etc.) in the language, include these in the investigation.

b. Implement, in a programming language of your choice, data structures and operations (empty, binding, lookup, enter and exit) for simple symbol tables.

c. In some programming languages, identifiers are case-insensitive so, e.g., size and SiZe refer to the same identifier. Describe how symbol tables can be made case-insensitive.