Assignment 1

Due date: 16-02-2019

1. Provide regular expressions to characterize the following lexical items:

   (a.) Identifiers: alphanumeric strings starting with an alphabet, and can also contain special character such as underscore.

   (b.) Numbers, which include integers, fixed point and floating point numbers. Leading zeros and redundant trailing zeros are disallowed. Real numbers can be represented in the E (exponential) format as well (e.g. $2.0E-2$ is valid and represents 0.002). Decisions whether $2.0E+2$ is valid are up to you.

   (c.) White space (sequences of blank, tab, newline)

   (d.) Arithmetic operators: $+, -, /, *$, div, mod

   (e.) Logical constants: true, false,

   (f.) Comparison operators: $<=, >=, <, >, =$

2. Which one of the following regular expressions represents the language: the set of all binary strings having two consecutive 0s and two consecutive 1s?

   (a) $(0+1)^*0011(0+1)^* + (0+1)^*1100(0+1)^*$

   (b) $(0+1)^*(00(0+1)^*11 + 11(0+1)^*00)(0+1)^*$

   (c) $(0+1)^*00(0+1)^* + (0+1)^*11(0+1)^*$

   (d) $00(0+1)^*11 + 11(0+1)^*00$

3. Convert the following NFAs to the equivalent DFAs in Figure 1.

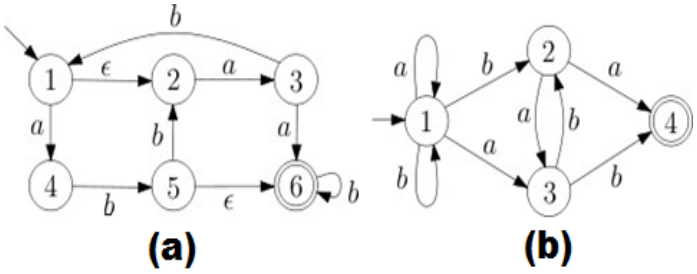

Figure 1: NFA

4. Minimize the following DFA in Figure 2 by using table filling method or Myhill-Nerode theorem.

5. Consider the alphabet $\sum = \{0, 1\}$, the null/empty string λ and the sets of strings X_0, X_1 and X_2 generated by the corresponding non-terminals of regular grammar. X_0, X_1 and X_2 are related as follows:
   $X_0 = 1X_1$

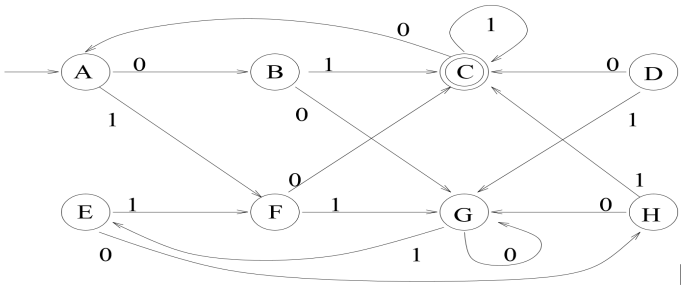

Figure 2: DFA

$X_1 = 0X_1 + 1X_2$
$X_2 = 0X_1 + \{\lambda\}$
Which one of the following choices precisely represents the strings in X_0?

(a) $10(0^* + (10)^*)1$

(b) $10(0^* + (10)^*)^*1$

(c) $1(0 + 10)^*1$

(d) $10(0 + 10^*)^*1 + 110(0 + 10)^*1$

6. The following CFG can parse all the lowercase roman numerals from $1 - 99$. The terminal symbols are $\{c, l, x, v, i\}$ and the initial symbol is S.
$S \rightarrow xTU|lX|X$
$T \rightarrow c|l$
$X \rightarrow xX|U$
$U \rightarrow iY|vI|I$
$Y \rightarrow x|v$
$I \rightarrow iI|\epsilon$

   (a.) Draw a parse tree for $62 : lxii$.

   (b.) Write semantic actions for each of the 14 rules in the grammar to calculate the decimal value of the input string. You can associate a synthesized attribute val to each of the non-terminals to store their value. The final value should be returned in $S.val$.

7. Calculate the first and follow functions for the given grammar
$S \rightarrow aBDh$
$B \rightarrow cC$
$C \rightarrow bC|\epsilon$
$D \rightarrow EF$
$E \rightarrow g|\epsilon$
$F \rightarrow f|\epsilon$

8. Compute the First and Follow sets as well as construct the parsing table for the following LL(1) grammars.
C → P F class id X Y
P → public $|\epsilon$

F → final |ϵ
X → extends id |ϵ
Y → implements I |ϵ
I → id J
J → , I |ϵ

9. (a.) Do left factoring in the following grammar
   $S \rightarrow aSSbS|aSaSb|abb|b$

   (b.) Remove left recursion of the following grammar
   A → B x y |x
   B → C D
   C → A |c
   D → d

10. Consider the grammar shown below:
    $S \rightarrow iEtSS'|a$
    $S' \rightarrow eS|\epsilon$
    $E \rightarrow b$
    In the predictive parse table, M, of this grammar, What will be the entries $M[S', e]$ and $M[S', \$]$?