

# CSE 401 - Homework Assignment #1

**Due: Wednesday, January 16, 2008 (at the beginning of class)**

0. Write regular expressions for each of the following:
  - a. All strings over the alphabet  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  that contain at least one  $\mathbf{a}$  and one  $\mathbf{b}$ .
  - b. All strings over the alphabet  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  where the first  $\mathbf{a}$  (if there is one) precedes the first  $\mathbf{b}$  (if there is one).
  - c. All strings over the alphabet  $\{\mathbf{0}, \mathbf{1}\}$  that contain an odd number of  $\mathbf{0}$ 's (a string must contain at least one  $\mathbf{0}$ ).
1. Convert the following regular expression to a NFA:  
 $\mathbf{a} ( (\mathbf{b} \mid \mathbf{a}^* \mathbf{c}) \mathbf{x} )^* \mid \mathbf{x}^* \mathbf{a}$
2. Convert the NFA into a DFA, following the algorithm from class. Be sure to label the NFA states and to label each of the DFA states with a set of NFA states.
3. The regular grammar specifying lexically correct programs for MiniJava is given as follows:  $\mathbf{Program} ::= (\mathbf{Token} \mid \mathbf{Whitespace})^*$ 
  - a. Modify this specification to require that all tokens be separated by whitespace, and optionally allow whitespace at the start and/or end of the program.
  - b. Why does this language-change remove the need for the longest-match meta-rule?
  - c. Do you think this would be a good language design change? (justify your answer)

Produce a hard-copy of your answers and turn them in by the start of class on the due date. Do these exercises individually, not with your project partner.