

CSE 401 Homework Assignment #1

Due: Monday, October 9, 2006

0. Write regular expressions for each of the following.
 - a. Strings over the alphabet $\{a, b, c\}$ where the first a precedes the first b .
 - b. Strings over the alphabet $\{a, b, c\}$ with an even number of a 's.
 - c. Strings over the alphabet $\{a, b, c\}$ that don't contain the contiguous substring baa .
1. In section A2.5.1 of *The C programming Language*, the lexical definition of integer constants is given as follows:

“An integer constant consisting of a sequence of digits is taken to be octal if it begins with 0 (digit zero), decimal otherwise. Octal constants do not contain the digits 8 or 9. A sequence of digits preceded by 0x or 0X (digit zero) is taken to be a hexadecimal integer. The hexadecimal digits include a or A through f or F with values 10 through 15.

“An integer constant may be suffixed by the letter u or U to specify that it is unsigned. It may also be suffixed by the letter l or L to specify that it is long.”

Write a regular expression (or collection of regular expressions) for this language of nonnegative integer constants in C.

2. Convert the following regular expression to a NFA:
 $a((bcd|a^*cd)x)^*|x^*a$
3.
 - a. Convert the following regular expression (where the alphabet is $a, b,$ and c) into an NFA, following the mechanical rules developed in class.
 $(a|b)(a|b)^*|(a|b)(a|b)^*c(a|b)(a|b)^*$
 - b. Convert this 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.
 - c. Describe the set of strings generated by this regular expression (e.g., an English description like the ones in question 0.)
4.
 - a. The regular grammar specifying lexically correct programs for MiniJava is given as follows: **Program** ::= **(Token|Whitespace)*** 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?

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.