CSE 420: Compiler Design

Assignment 1: Regular Expression to Direct DFA Construction

Submission Deadline: Monday, Nov 11 (In class)

In this assignment, you have to convert a regular expression into a DFA using the direct RE to DFA construction process explained in the reference textbook. This is your first, self-study theory assignment. To understand how the direct RE-to-DFA construction process works, you have to carefully read Section 3.9.1 to Section 3.9.5 of the reference textbook. There are four steps in the direct RE-to-DFA construction process. First, you have to augment the regular expression with an end marker and draw a regular expression tree; then you compute NULLABLE, FIRSTPOS, and LASTPOS for each node of the expression tree; then you compute FOLLOWPOS for each non-empty leaf of the expression tree, finally, you construct a table representing your DFA from the FOLLOWPOS information. You have to show each step of the DFA construction process in your hard-copy submission. Note that your submission must be a handwritten version, a typed submission will get a 0 out of 10. Finally select your regular expression, using the following process:

1. Take your 8-digit long student ID.
2. The first letter of your RE is the letter you get from the English alphabet by doing a first-digit % 3 of your student ID. (Note that the letters of the alphabet are indexed here from 0 to 25).
3. Similarly, the second letter is, second digit % 3.
4. Assume the third-digit % 3 is ‘a’ then you add (a|ε) to the RE. Note that ε means epsilon.
5. Then add seventh-digit % 3 to the RE. Then add a Kleen’s closure *.
6. Finally add eight-digit % 3 to the RE.

An example: Suppose your student ID is 20130165. Then

2%3 = 2, letter is c

0%3 = 0, letter is a

1%3 = 1, letter is b

6%3 = 0, letter is a

5%3 = 2, letter is c

Then the regular expression will be ca(b|ε)a*c

Late Submission Policy: You will miss 2 points for each day of delay after the deadline. So, there is no point submitting the assignment after 5 days as you will get a zero regardless of your answer.