CSE340 Fall 2018 - Homework 1

Due: Thursday September 6, by 11:59 PM on Blackboard

All submissions should be typed. Exception can only be made for drawing parse trees, which can be hand drawn and scanned in the submitted document.

You should write your solutions in order. Solution to problem 1 should be first, then solution to problem 2, then solution to problem 3 and finally solution to problem 4.

**Problem 1.** Consider the following regular expressions

R0 = 1 | 2 | 3

R1 = 1 | 2 | 4 | 8

R2 = (a | b) (a\* | b\*) (a | b)

R3 = (a\* | b\*) R1\* (ab)\*

R4 = ab R3\* (a | b)\*

R5 = R3\* aaa R2\*

In the expressions, the dot operator is omitted and some parentheses are omitted, in which case the Kleene star operator (\*) has the highest precedence, followed by the dot operator (.), followed by the or operator (|).

Let getToken() be a function that returns the next token in the input. If we call it repeatedly it will return one token after another. When all the input is consumed, getToken() returns EOF (end of file). Assume that longest prefix-matching rule is used by getToken() and ties are broken in favor of the regular expression listed first.

1. Give an example of input for which calling getToken() twice returns R0 first then EOF
2. Give an example of input for which calling getToken() twice returns R1 first then EOF
3. Give an example of input for which calling getToken() twice returns R2 first then EOF
4. Give an example of input for which calling getToken() twice returns R3 first then EOF
5. Give an example of input for which calling getToken() twice returns R4 first then EOF
6. Give an example of input for which calling getToken() twice returns R5 first then EOF
7. If getToken() is called repeatedly on the following input, what is the sequence of tokens and lexemes returned? In your work, show step by step the Matched, Potential (Viable), and Maximal tokens.

aaa1ba1ba1daa1ab

**Problem 2.** Let R1 and R2 be two regular expressions over the alphabet {a , b}

1. Is it always the case that L(((R1).(R2))\*) = L(((R1)|(R2))\*)?
2. Is it always the case that L(((R1)\*.(R2)\*)\*) = L(((R1)|(R2))\*)?
3. Is it always the case that L((R1\*)\*) = L((R1)\*)?

In each case explain why the statement is true or provide a counter example.

**Problem 3.** Consider the grammar

S → Y X Y

X → a X | Y

Y → b b Y | X | ε

Draw a parse tree for input string abbaabb

**Problem 4.** Consider the grammar

S → A

S → a

A → a b S | a b S d S

1. What are the non-terminals?
2. What is the start symbol?
3. What are the terminals?
4. Show that this grammar is ambiguous (this is not an easy question and would require some effort to find a input that has two parse trees)