CS164 Programming Languages and Compilers **Spring 2024**

Written Assignment 1

Assigned: January 24 **Due:** February 7 at 11:59 pm

Instructions: This assignment asks you to prepare written answers to questions on lexical analysis, regular expressions, and finite automata. Each of the questions has a short answer. You may discuss this assignment with other students and work on the problems together. However, your write-up should be your own individual work.

Please write your name, email address, and discussion section on your homework. ***Please start each question on a new page. All written assignments must be submitted as a PDF via Gradescope: https://gradescope.com.*** Instructions for how to submit assignments to Gradescope can be found at the following links: `https://gradescope.com/get_started#student-submission`

1. Consider the following languages of binary numbers over the alphabet $\Sigma = \{0, 1\}$.

   - L_1: *All binary numbers where the last digit is a 1* (e.g. $1, 011, 111, 1011 \in L_1$)
   - L_2: *All binary numbers divisible by 4*
   - L_3: *All binary numbers divisible by 3*
   - L_4: *All binary numbers that contain exactly 2 0's or no 1's*

   Give a deterministic finite automaton (DFA) for all the languages above. (Note: Empty strings are not binary numbers.)

2. Consider the regular expression $R = (ab)^* \mid (bb \mid aba)^*$, note that language $L(R)$ is over the alphabet $\Sigma = \{a, b\}$

   - Construct an ϵ-NFA for the language $L(R)$.
   - Convert the above NFA to DFA.

   (**Hint:** use approach describe in the lecture : ϵ-NFA $\rightarrow$ NFA (label states) $\rightarrow$ DFA).

3. Let $\Sigma_m = \{a_1, \ldots, a_m\}$ be an alphabet containing m elements, for some integer $m \geq 1$. Let L_m be the following language that includes all strings in which at least one of the characters occurs an even number of times, i.e.

   All strings in which a_i occurs an even number of times for some i, where $1 \leq i \leq m$

   Construct a DFA for the language L_3. Also construct an NFA for the language L_4.

4. Determine whether or not the following languages are regular. Explain why in one or two sentences.

   - L_1: *All strings over the alphabet* $\{a, b\}$ *where there are at least as many a's as there are b's.*
   - L_2: *All strings over the alphabet* $\{a, b\}$ *that are palindromes (same string when reversed).*
   - L_3: *All words in the Oxford English dictionary.* (**Hint:** assume dictionary has finite number of words).

5. Let $\Sigma = \{a, b\}$ be the alphabet for the language $L = \{waw^R \mid w, w^R \in \{a, b\}^*, \text{ and } w \text{ has even length}\}$, where w^R is the reverse of w.

   Write a context free grammar for the language L.

6. Consider the following grammar:

$$S \to [\ S\ S\ ]$$
$$S \to a$$
$$S \to \varepsilon$$

Show that this grammar is ambiguous by finding a string that can be parsed in at least three different ways. Draw three different parse trees for this string, and write down the left-most derivation for each of the three trees.

7. Give context free grammars for the following languages. Your grammars should not be unnecessarily complex. For each grammar, briefly explain why your grammar accepts precisely the specified language.

   (a) $L = \{x^i y^j : 0 \leq j \leq i\}$, where for example $x^5 y^2 = xxxxxyy$.

   (b) $L = \{a^i b^j c^k : i \geq 0,\ j \geq 0,\ \text{and } i + j \leq k\}$.

   (c) $L = \{w^R \# w \mid w \in \{0,1\}^* \text{ and } w \text{ as a binary number is divisible by } 3\}$. For this problem, if w is empty then its value as a binary number is 0.