CS323 Assignment 1

1 Requirements

You are expected to complete all required homework exercises and encouraged to complete the optional ones (if there are). For submission, please put all your answers in a single PDF file and submit it via the assignment channel on BlackBoard. The name of the file should follow the format **"studentID_A#"** (e.g., 30003554_A1). **The submission deadline is 10:00 PM, October 2, 2023.** Late submissions are allowed within three days after the deadline (grace period). If you submit your assignment during the grace period, your score will be 80% of the score you could get if the submission was made in time. Assignments submitted after the grace period will not be graded.

2 Required Exercises (100 points)

Exercise 1: Java programs may contain lexical errors. Please give at least two types of possible lexical errors in Java programs and provide code snippets as examples when possible. [10 points]

Exercise 2: Given a string s, can you find a string x that is both a prefix and a suffix of s? Can you further find a string y that is both a proper prefix and a proper suffix of s? If yes, please provide an example. Otherwise, please explain the reason. [10 points]

Exercise 3: In a string of length n ($n > 0$), how many of the following are there? For simplicity, we assume that the string contains n different characters. Besides giving the final answers, please also explain how you derive the answers.

1. Substrings of length m ($0 < m \leq n$) [15 points]
2. Subsequences [15 points]

Exercise 4: Write a regular definition as well as a regular expression to represent all strings of valid telephone numbers in Shenzhen. A valid telephone number contains the country code (86), a hyphen, the area code 755, another hyphen, and eight digits where the first one cannot be zero (e.g., 86-755-88015159). [10 points]

Exercise 5: Given an alphabet $\Sigma = \{0, 1\}$, are the following two regular languages equivalent? Besides saying yes or no, please also prove your answer. [20 points]

1. $L_1 = L((0^*1^*)^*)$
2. $L_2 = L((0|1)^*)$

Exercise 6: Consider the regular expression $ba^*|ab^*$. Please provide a state transition diagram that can recognize the strings in the corresponding regular language. Can the transition diagram recognize the string $baab$? If yes, please give the sequence of state transitions. Otherwise, please explain the reason. [20 points]