CS323 Assignment 1

1   Requirements

You are expected to complete all required homework exercises and encouraged to complete the optional ones. For submission, please put all your answers in a single PDF file and submit it via the assignment channel on SAKAI. The name of the file should follow the format **"studentID_A#"** (e.g., 30003554_A1). **The submission deadline is 11:55 PM, September 27, 2022.** Late submissions are allowed within one week 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. Assignment submitted after the grace period will not be graded, meaning that you will get a zero for the assignment.

2   Required Exercises (100 points)

Exercise 1: When a `C` compiler compiles the following statement, how many tokens will it generate? [5 points]

`int a3 = a3 * 3;`

Exercise 2: 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.

1. Prefixes [5 points]

2. Proper prefixes [5 points]

3. Prefixes of length m $(0 < m \leq n)$ [5 points]

4. Suffixes of length m $(0 < m \leq n)$ [5 points]

5. Proper prefixes of length m $(0 < m \leq n)$ [10 points]

6. Substrings [10 points]

7. Subsequences [10 points]

Exercise 3: Describe the regular languages denoted by the following regular expressions:

1. `((ε|a)*b*)*` [5 points]

2. `(a|b)*a(a|b)(a|b)` [5 points]

3. `a*ba*ba*ba*` [5 points]

Exercise 4: Write regular definitions or regular expressions for the following languages.

1. All strings representing 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]

2. All strings of a's and b's that start with a and end with b. [10 points]

3. All strings of lowercase letters that 1) contain the five vowels (i.e., a, e, i, o, u) and 2) the vowels appear in order. For example, abaeeiou is such a string (we allow that a vowel appears multiple times before its next one appears) but aeaiou is not as the second 'a' appears after 'e'. [10 points]

3 Optional Exercises (10 bonus points)

Exercise 1: Given an alphabet $\Sigma = \{a, b\}$, are the following two regular languages equivalent? Please also prove your answer.

1. $L_1 = L((a^*b^*)^*)$

2. $L_2 = L((a|b)^*)$