Assignment 1

1 Requirements

You are expected to complete all required exercises and encouraged to complete the optional ones. For submission, please put all your answers in a PDF file. The name of the file should follow the format "studentID assignmentNumber" (e.g., 2021111054_assignment1). The submission deadline is 20230924. Late submissions are still allowed within one week after the deadline (grace period). If you submit your assignment during the grace period, your score will be half of the score you could get if the submission was made before the deadline. 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 = a * 3;
```

Exercise 2

In a string of length n $(n > 0)$, how many of the following are there?

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 languages denoted by the following regular expressions:

1. $((\epsilon|a)^*b^*)^*$ [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 010, another hyphen, and eight digits where the first one cannot be zero (e.g., 86-010-62282045). [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 contain the five vowels in order. [10 points]

3 Optional Exercises (10 bonus points)

Exercise 1

Suppose we have a alphabet $\Sigma = a,b,c$, write regular definitions to describe all strings over Σ without repeated letters. [Hint: You may draw an NFA for the language and convert the NFA to regular definitions.]