Tutorial 7

Exercise 1: For each of the following languages, give an example of 5 words belonging to the language, and an example of 5 words that do not belong to the language.

a) $L_1 = \{w \in \{0,1\}^* \mid$ the length of word w is less than 5$\}$
b) $L_2 = \{w \in \{a,b\}^* \mid$ the number of occurrences of symbol b in word w is even$\}$
c) $L_3 = \{w \in \{0,1\}^* \mid$ in w is every 0 (directly) followed by $1\}$
d) $L_4 = \{w \in \{0,1\}^* \mid w$ begins and ends with the same symbol$\}$
e) $L_5 = \{w \in \{a,b\}^* \mid w$ contains as a subword the sequence $abb\}$

Exercise 2: Let us assume $\Sigma = \{a,b\}$ and $n \in \mathbb{N}$.

a) How many words in Σ^* are of length n?

b) How many words in Σ^* are of length at most n?

Exercise 3: Consider the following languages:

$L_1 = \{w \in \{0,1\}^* \mid$ in w is every 0 (directly) followed by $1\}$
$L_2 = \{w \in \{0,1\}^* \mid w = w^R\}$

a) Enumerate the first 5 words of each of languages L_1, L_2 (the smallest words with respect to order $<_L$).

b) Enumerate the first 5 words of each of languages $\overline{L_1}$, $\overline{L_2}$.

c) Enumerate the first 5 words of language $L_1 \cap L_2$.

d) Enumerate the first 5 words of language $L_1 \cup L_2$.

Exercise 4: Consider languages over $\{a,b\}$. Write down all the words in the concatenation of $L_1 = \{\varepsilon, abb, bba\}$ and $L_2 = \{a, b, abba\}$.

Exercise 5: Consider languages over the alphabet $\{0,1\}$. Write down all words in the concatenation

$$\{0, 001, 111\} \cdot \{\varepsilon, 01, 0101\}$$

Exercise 6: Consider languages over the alphabet $\{0,1\}$. Describe the language of all words in the iteration $\{00, 111\}^*$ and write the first 10 words of the language.

Exercise 7: Consider the following languages:

$L_1 = \{w \in \{0,1\}^* \mid |w|_1 \leq 1\}$
$L_2 = \{w \in \{0,1\}^* \mid w = w^R\}$

Describe the words in the language $L_1 \cap L_2$.

Exercise 8: Write regular expressions for the following languages:

a) The language $\{ab, ba, abb, bab, abbb, babb\}$
b) The language over alphabet $\{a, b, c\}$ containing exactly those words that contain subword abb.
c) The language over alphabet $\{a, b, c\}$ containing exactly those words that start with prefix bca or end with suffix $ccab$.
d) The language $\{w \in \{0,1\}^* \mid |w|_0 \bmod 2 = 0\}$.
e) The language $\{w \in \{0,1\}^* \mid |w|_0 \bmod 3 = 1\}$.
f) The language $\{w \in \{0,1\}^* \mid w$ contains subwords 010 and $111\}$
g) The language $\{w \in \{a,b\}^* \mid w$ contains subword bab or $|w|_b \leq 3\}$
h) The language $\{w \in \{a,b\}^* \mid w$ contains subword bab and $|w|_b \leq 3\}$
i) The language of all words over $\{a, b, c\}$ that contain no two consecutive a's.

Exercise 9: Let us have two languages K and L described by the regular expressions

$$L_1 = \mathcal{L}(\, 0^*1^*0^*1^*0^* \,), \qquad L_2 = \mathcal{L}(\, (01 + 10)^* \,).$$

a) What are the shortest and the longest words in the intersection $L_1 \cap L_2$?

b) Why none of the languages L_1 and L_2 is a subset of the other?

c) What is the shortest word that does not belong to the union $L_1 \cup L_2$? Is it unambiguous?

Exercise 10: Let us say that we would like to devise a syntax for representation of simple arithmetic expressions by words over alphabet

$$\Sigma = \{\texttt{A}, \texttt{B}, \ldots, \texttt{Z}, \texttt{a}, \texttt{b}, \ldots, \texttt{z}, \texttt{0}, \texttt{1}, \ldots, \texttt{9}, \texttt{.}, \texttt{+}, \texttt{-}, \texttt{*}, \texttt{/}, \texttt{(}, \texttt{)}\}.$$

a) Propose how identifiers will look like, and deribe them using a regular expression.

b) Propose how number constants will look like, and describe them using a regular expression.

Remark: Allow the number constants that would represent integers, e.g., `129` or `0`, and also floating-point number constants, e.g., `3.14`, `-1e10`, or `4.2E-23`. Consider also the possibility of representing number constants in other number systems except the decimal number system (e.g., hexadecimal, octal, binary).