4 LAB 0: GRAMMARS AND FORMAL LANGUAGES

4.1 Theory assignment

These assignments are **optional** for those who need to revise their theory knowledge of regular expressions and formal languages.

The lab is intended to make you practice using the basic concepts used in descriptions of formal languages.

1. Given the alphabet A = { 1, 2, 3 } and the strings x = 2, y = 13, z = 323:

    a. What are the strings xz, zyx, z^2, x^7 and what are their lengths?

    b. What are A^1, A^2, A^0

    c. Describe A^* and A^+

2. Given the following grammar for arithmetic expressions G(<exp>):

                        <exp>   ::= <term> | <exp> + <term> | <exp> - <term>

                        <term>   ::= <factor> | <term> * <factor> | <term> / <factor>

                        <factor> ::= (<exp>) | <ident>

                        <ident>   ::= A | B | C ... | Z

    a. Find derivations and draw a parse tree for the statements: A*B-C, A*(B-C), A/B/C and -A*B

    b. State the handle in statement form <term> * <factor> + B/C

    c. State a canonic derivation of one of the statements in example 2a.

    d. Why is A * <exp> * D not in statement form?

    e. What are V, Σ and N in this grammar?

    f. Describe $Σ^+$ and L(G). State some string in $Σ^+$ which is not in L(G).

3. Using the grammar in exercise 2, draw the parse tree for the statement

                A - <term> * B + (<term>)

    Use this to explain why there is no canonic derivation of the statement.

4. Construct a context-free grammar for even integers which may not begin with zeros (this also applies to a single zero).

5. Provide a regular expression for strings consisting of binary numbers separated by the operators + or -. Example:

                10+1001-01

                111001

    Draw the respective state diagram, too.

6. Given the alphabet A = { a, b, c, d, e, f, g, h, i }:

    a. Construct a context-free grammar for words of one syllable (i.e. words containing exactly one vowel).

b. Construct a context-free grammar for multi-syllable words.

7. Given the alphabet A = { a, b, c }:

Write regular expressions which describe

a. all strings starting with a sequence of zero or several a's followed by an arbitrary number of b's and c's mixed together. The strings finish with another sequence of a's. (Example: aaabbcbca, bbca, aa).

b. $a^n b^n c^n$, can this expression be described using a regular expression? If so, write down the regular expression. If not, motivate why.

8. Describe the languages (sets) that the following grammars generate:
<S> is the start symbol.

a.

<S> ::= <A> <A>

<A> ::= 1 <A> 0 | 10

b.

<S> ::= 1 <S> 0 | <B>

<B> ::= 0 <B> 1 | ε

c.

<S> ::= 1 <A> | <B> 0

<A> ::= 1 <A> | <C>

<B> ::= <B> 0 | <C>

<C> ::= 1 <C> 0 | ε