Exercise session 5

1. How could you use regular expressions and finite automata in practice? Give at least three applications!

2. Are the following languages regular? Justify your answer carefully!

   (a) $\{w|w$ is a character string of length 3 consisting of a's and b's$\}$

   (b) $\{ww^*|w \in \{a,b\}^*\}$

   (c) $\{w^*|$there are as many 1's as 0's in w

   (d) $\{w|$there are twice as many 0's as 1's in w$\}$

3. What happens if you try to prove an actually regular language as non-regular by the Pumping Lemma? Consider for example languages $L_1 = \emptyset$, $L_2 = \{aa, bb\}$, $L_3 = \{ab^*a^*b\}$.

4. What is wrong in the following Pumping lemma proofs?

   a) Let $L = \{(aa)^i(bb)^j|i,j \geq 0\}$.
   Claim: L is nonregular.
   Proof: Let $x = (aa)^n(bb)^k = a^{2n}b^{2n}$, $|x| = 4n$. Now x can be divided into parts uvw only in one way: $u = a^i$, $v = a^j$, $j \geq 1$, $w = a^{2n-i-j}b^{2n}$. Now $uv^0w = a^{2n-j}b^{2n} \notin L$, thus L is non-regular.

   b) Let $L = \{c^ra^kb^k|r \geq 1, k \geq 0\} \cup \{a^kb^l|k,l \geq 0\}$.
   Claim: L is regular.
   Proof: $L = L_1 \cup L_2$, in which $L_1 = \{c^ra^kb^k|r \geq 1, k \geq 0\}$ and $L_2 = \{a^kb^l|k,l \geq 0\}$. For all $x \in L$ either $x \in L1$ or $x \in L2$. Let's consider both cases:
   1) If $x \in L_1$, let's select $x = c^na^kb^k$. $|x| = n + 2k > n$. x can be divided into parts uvw only in one way:
   $u = c^i$, $v = c^j$, $j \geq 1$, $w = c^{n-i-j}a^kb^k$. Now $uv^kw \in L$ for all $k = 0,1,2,...$, i.e. we can always pump x.
   2) If $x \in L_2$, let's select $x = a^nb^l$, $|x| = n + l$. x can be divided into parts in only one way $u = a^i$, $v = a^j$, $j \geq 1$, $w = a^{n-i-j}b^l$. $uv^kw \in L$ for all $k = 0,1,2,....$
   $\Rightarrow L$ is regular.

5. Prove with the Pumping lemma that the following languages are non-regular, and give the context-free grammars which describe the languages.

(a) $\{a^n b^n c^k | n, k = 0, 1, ...\}$

(b) $\{a^n b^k c^k | n, k = 0, 1, ...\}$

(c) $\{a^n b^n a^m b^m | n, k = 0, 1, ...\}$

6. Let's denote by w^R the string w written backwards (i.e. if $w = a_1 a_2 \dots a_n$, then $w^R = a_n \dots a_2 a_1$). The string is a palindrome if $w = w^R$ (e.g. "madamimadam"). Consider the palindrome language $L_{pal} = \{ww^R | w \in \{a, b\}^*\}$.

   Prove with the Pumping Lemma that L_{pal} is non-regular!

7. Let's have a context-free grammar $G = (V, \Sigma, P, S)$, $V = \{A, B, C, D, S\} \cup \Sigma$, in which $\Sigma = \{Jim, big, green, cheese, ate\}$, and rules P are:
   $A \rightarrow B|CA$
   $S \rightarrow ADA$
   $C \rightarrow big|green$
   $B \rightarrow cheese|Jim$
   $D \rightarrow ate$

   What kind of strings belong to the language described by G? Give some exaples! Do the following sentences belong to the language $L(G)$: "big big green cheese", "ate Jim big cheese"?

8. (a) Construct a finite automaton, which recognizes the language described by the following grammar. What is the corresponding regular expression?
   $S \rightarrow aA|bB$
   $A \rightarrow aS|bA$
   $B \rightarrow bB|\epsilon$

   (b) Give a right-linear grammar that describes the language
   $L = \{w \in \{a, b\}^* | w \text{ does not contain substring } abaa\}$

9. All words of the language of outerspace aliens follows the Blurbs normalform. Blurb is a Whoozit, followed by one or more Whatzit. Whoozit is letter 'x', which can be followed by any number of letters 'y' (also zero). Whatzit is letter 'q', which is followed by either letter 'z' or 'd', followed by Whoozit.

   Give the context-free grammar, which describes the Blurbs language. Can you now give the corresponding regular expression? (Hint: construct first the automaton from the grammar!)

10. For regular expression holds the following rule::
If $r = rs \bigcup t$, then $r = ts^*$, when $\epsilon \notin L(s)$.
Show, why this rule holds with the help of context-free grammars!

More challenging:

11. We know that in general the problem $REG(L)$ (i.e. is the given language L regular or not) is unsolvable. What is the reason for that? Could you still invent a program, which would help people to prove by the Pumping Lemma that the given language is nonregular?

12. The stronger version of the Pumping Lemma gives both sufficient and necessary conditions for the regularity of a language:

    Pumping Lemma 2: The language $A \in \Sigma^*$ is regular if and only if there is a constant $n \geq 1$ such that for all $x \in \Sigma^*$, if $|x| \geq n$ then there exists u, v, w such that $x = uvw$ and $|v| \geq 1$, and for all $i \geq 0$ and all $yin\Sigma^*$ $xy \in A$ if and only if $uv^iwy \in A$.

    Based on this, could you make a program, which decides for any given language A, if it is regular or not?