T–79.1001 **Autumn 2007**
Introduction to Theoretical Computer Science T
Tutorial 7, 1 to 6 November
Problems

Homework problems:

1. Show, using the pumping lemma for regular languages, that the language consisting of even-length palindromes,

$$\{ww^R \mid w \in \{a,b\}^*\}$$

is not regular.

2. Convert the following grammar into Chomsky normal form:

$$\begin{aligned} S &\rightarrow AB \mid c \\ A &\rightarrow T \mid aA \\ B &\rightarrow TT \mid \varepsilon \\ T &\rightarrow bS \end{aligned}$$

3. Determine, using the CYK algorithm ("dynamic programming method", Sipser p. 241, Lewis & Papadimitriou p. 155), whether the strings *abba*, *bbaa* and *bbaab* are generated by the grammar

$$\begin{aligned} S &\rightarrow AB \mid BA \mid a \mid b \\ A &\rightarrow BA \mid a \\ B &\rightarrow AB \mid b \end{aligned}$$

In the positive cases, give also the respective parse trees.

Demonstration problems:

4. *Pattern expressions* are a generalisation of regular expression used e.g. in some text editing tools of UN*X-type operating systems. In addition to the usual regular expression constructs, a pattern expression may contain string variables, inducing the constraint that any two appearances of the same variable must correspond to the same substring. Thus e.g. aXb^*Xa and $aX(a \cup b)^*YX(a \cup b)^*Ya$ are pattern expressions over the alphabet $\{a,b\}$. The first one of these describes the language $\{awb^nwa \mid w \in \{a,b\}^*,\ n \geq 0\}$. Prove that pattern expressions are a proper generalisation of regular expressions, i.e. that pattern expressions can be used to describe also some nonregular languages.

5. Prove that the language $\{w \in \{a,b\}^* \mid w$ contains equally many a's and b's$\}$ is not regular, and design a context-free grammar generating it.

6. Design an algorithm for testing whether a given a context-free grammar $G = (V, \Sigma, P, S)$, generates a nonempty language, i.e. whether any terminal string $x \in \Sigma^*$ can be derived from the start symbol S.