T-79.148 **Spring 2003**
Introduction to Theoretical Computer Science
Tutorial 6, 24–26 February
Problems

Homework problems:

1. Consider the following context-free grammars:

$$\text{(a)} \quad \begin{array}{rcl} A & \rightarrow & aAcc \mid B \\ B & \rightarrow & bBc \mid \varepsilon \end{array} \qquad\qquad \text{(b)} \quad S \rightarrow aSb \mid SS \mid \varepsilon$$

Give a derivation for the sentence *abccc* according to grammar (a), and a derivation for the sentence *aababbab* according to grammar (b). Describe the language generated by each grammar verbally as simply as you can.

2. A *palindrome* is a string w such that $w = w^R$. (E.g. "MADAMIMADAM", "ABLEWASIEREISAWELBA," cf. `http://www.palindromes.org/`.) Consider the set of palindromes over the alphabet $\{a, b\}$:

$$\mathrm{PAL} = \{w \in \{a,b\}^* \mid w = w^R\}.$$

   (a) Prove that this language is not regular.
   (b) Design a context-free grammar generating the language.

3. Design a context-free grammar corresponding to the following XML/DTD description:

```
<!DOCTYPE Book [
   <!ELEMENT Book (Title, Chapter+)>
   <!ATTLIST Book Author CDATA #REQUIRED>
   <!ELEMENT Title (#PCDATA)>
   <!ELEMENT Chapter (#PCDATA)>
   <!ATTLIST Chapter id ID #REQUIRED>
]>
```

(This example appears on the WWW page `http://www.rpbourret.com/xml/xmldtd.htm`, which also contains further explanations regarding its interpretation and the DTD notation in general. Context-free grammars and the DTD notation do not correspond to each other quite one-to-one, so solving the problem may require some independent interpretation of the conventions used.)

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.

PLEASE TURN OVER

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

6. Design a context-free grammar describing the syntax of simple "programs" of the following form: a program consists of nested `for` loops, compound statements enclosed by `begin-end` pairs and elementary operations `a`. Thus, a "program" in this language looks something like this:

```
a;
for 3 times do
begin
   for 5 times do a;
   a; a
end.
```

   For simplicity, you may assume that the loop counters are always integer constants in the range 0,..., 9.