

Tutorial 1: Problems

CSC 467 Compilers and Interpreters
Fall Semester, 2005

1. (**Thompson's construction**) Page 149, problem 3.16.
2. (**Subset construction**) Page 150, problem 3.17. Do parts a) and c) only.
3. (**Regular Expressions**) Give regular definitions for the following languages:
 - (a) All strings of letters that contain the five vowels in order.
 - (b) All strings of letters in which the letters are in ascending lexicographic order.
 - (c) All strings of 0's and 1's that do not contain the substring 011.
 - (d) All strings of 0's and 1's that do not contain the subsequence 011.
 - (e) All strings of 0's and 1's with exactly one appearance of 000.
4. (**Regular Expressions cont'd**) Give the simplest regular expression that describes the same language as the following regular expressions:
 - (a) $(a|b)^*b^+b^*$
 - (b) $(a^*(ab)^*(abc)^+abc$
 - (c) $((p^+q^*)^*|(q^+p^*)^*)^*r$
5. (**Context Free Grammars**) Give a Context Free Grammar for each of the following languages:
 - (a) $\{a^n b^m, 0 \leq n \leq m \leq 2n\}$
 - (b) $\{w \in (a|b)^*, w = w^R\}$ where w^R is the reverse of w .
 - (c) Strings of a 's and b 's where the number of b 's is double the number of a 's.
6. (**Left-Recursion and Left-Factoring**) Consider the following Pascal-like grammar where terminals are in boldface:

$$\begin{aligned}
 \langle var_decl \rangle &\rightarrow \mathbf{var} \langle decl_list \rangle \\
 \langle decl_list \rangle &\rightarrow \langle decl \rangle ; \langle decl_list \rangle \mid \langle decl \rangle \\
 \langle decl \rangle &\rightarrow \langle id_list \rangle : \mathbf{id} \\
 \langle id_list \rangle &\rightarrow \langle id_list \rangle , \mathbf{id} \mid \mathbf{id}
 \end{aligned}$$

- (a) Eliminate left-recursion and left-factor the grammar.
- (b) Find the **FIRST** and **FOLLOW** sets for all the non-terminal symbols of the new grammar of part (a).

Grammar Symbol	FIRST set	FOLLOW set