First exercise sheet for the lecture

XML and Programming Languages

Dr. Janis Voigtländer

Summer Term 2008

Exercise 1
Which of the following regular expressions are deterministic?

1. $(a|b)^*a$
2. $(a|b)^*a(a|b)$
3. $c(a|b(\epsilon|cc))^*$
4. $a^*(bca^*)^*$
5. $(a|b)^*(ac|bd)$ ◇

Exercise 2
Let d be the DTD (r, P) with $P = \{r \to a^*,\ a \to b|c,\ b \to \epsilon,\ c \to \epsilon\}$. Is $\mathcal{L}(d)$ regular? ◇

Exercise 3
Try to give a general construction delivering for every non-recursive DTD d an FSA recognizing $\mathcal{L}(d)$. ◇

Exercise 4

1. Recall the DTD $(r, \{r \to a,\ a \to a|\epsilon\})$ from the lecture, for which $\mathcal{L}(d)$ was not regular. Assume that a validator will only get well-balanced input strings. Is it then possible to detect, among those only, the string representations of tree documents valid with respect to d?
2. What about the DTD $(r, \{r \to a|\epsilon,\ a \to b,\ b \to a|\epsilon\})$?
3. Or the DTD $(a, \{a \to b^*,\ b \to a^*\})$?
4. Or the DTD $(r, \{r \to aa,\ a \to a|\epsilon\})$?

5. Or the DTD $(a, \{a \rightarrow ab|ca|\epsilon,\ b \rightarrow \epsilon,\ c \rightarrow \epsilon\})$?
6. Try to formalize a notion of *weak validation* capturing the above idea. $\diamond$