

Second exercise sheet for the lecture

XML and Programming Languages

Dr. Janis Voigtländer

Summer Term 2009

Exercise 4

Consider the DTD $d = (r, \{r \rightarrow a^*, a \rightarrow bc, b \rightarrow c|\epsilon, c \rightarrow \epsilon\})$. Using the construction from the lecture, give an FSA recognizing exactly $\mathcal{L}(d)$. \diamond

Exercise 5

1. Recall the DTD $(r, \{r \rightarrow a, a \rightarrow a|\epsilon\})$ 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 $(r, \{r \rightarrow a|\epsilon, a \rightarrow b, b \rightarrow a|\epsilon\})$?
3. Or the DTD $(a, \{a \rightarrow b^*, b \rightarrow a^*\})$?
4. Or the DTD $(r, \{r \rightarrow aa, a \rightarrow a|\epsilon\})$?
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

Exercise 6

For those DTDs d from Exercise 5 for which weak validation does not work, give context free grammars that generate exactly $\mathcal{L}(d)$. \diamond