Technische Universität München **Winter term 2011/12**
I7
Prof. J. Esparza / J. Křetínský **10.11.2011**

Automata and Formal Languages – Homework 5

Due 17.11.2011.

Exercise 5.1

In the lecture, you have seen that we can save on space using the *lazy* DFAs. However, this does not come for free. There is a space vs. running-time trade-off because of extra steps with head not moving in the case with lazy DFAs.

Find a word w and a pattern p such that the run of **eagerDFA**(p) on w takes at most n steps and the run of **lazyDFA**(p) takes at least $2n - 1$ steps.

Hint: a simple pattern of the form a^k is sufficient.

Exercise 5.2

Design an algorithm that solves the following problem for a finite alphabet Σ. Discuss the complexity of your solution.

- *Given*: $w \in \Sigma^*$ and a regular expression r over Σ.
- *Find*: A shortest prefix $w_1 \in \Sigma^*$ of w such that there exists a prefix $w_1 w_2$ of w and $w_2 \in \mathcal{L}(r)$.

Exercise 5.3

The constructed **eagerDFA** for a pattern p has a state for each prefix of p, hence $|p| + 1$ states altogether. Consider the simple NFA $\mathcal{N}$ for $\Sigma^*.p$ (with a self loop under Σ in the intial state and a transition from each prefix q of length $0 \leq |q| < |p|$ labelled by $p[|q| + 1]$ to $q.p[|q| + 1]$). How many states does NFAtoDFA($\mathcal{N}$) have? How long does it take to construct it?

Exercise 5.4

As seen in the lecture, when applying the post, pre or join operations to transducers the underlying projection operation might yield an automaton which does not accept all possible encodings anymore. You have seen how to fix this problem in the case that the representations are obtained by padding on the right, as for instance in the lsbf-representation of natural numbers where all representations of a number $n \in \mathbb{N}$ are obtained by adding 0s. In the case of the msbf-encoding, the padding does not occur on the right, but on the left. Hence, the procedure given in the lecture cannot be applied anymore.

- Give an algorithm for calculating the "pad-closure" of a transducer when using the msbf-encoding of natural numbers.