Assignment 2
for Formal Languages and Automata Theory
TDDD85

Deadline: Monday week 21 (19 May 2014)

For all the problems below it is not sufficient just to give a solution. **Justify your answers**. In the final exam unexplained answers will be granted 0 points.

The solutions should be handed in to Jonas Wallgren or Johannes Schmidt – on paper using the "från studenter till lärare" postbox in front of "Java Cafe" in building B (or at a lecture or a tutorial), or in a file in an email attachment (please only plain text, or ps or pdf).

Your solutions can be written in Swedish or English.

1. Choose an alphabet Σ with more than 1 elements and construct a context-free grammar for the language

$$L = \{\, ew_1ew_2ew_3e \mid e \in \Sigma,\ w_1, w_2, w_3 \in \Sigma^*,\ w_1 = w_2^R \,\}.$$

   (x^R denotes the string x reversed.)

2. Let $L^R = \{\, x^R \mid x \in L \,\}$. Which of the pairs of languages below differ (for some L, or L_1, L_2)?

$$\begin{array}{lll} LL & \text{and} & \{xx \mid x \in L\} \\ L_1^R L_2^R & \text{and} & (L_1 L_2)^R \\ L_1^R L_2^R & \text{and} & (L_2 L_1)^R \\ L^* & \text{and} & L^* L \end{array}$$

   For the pairs that differ, can you find special cases when they are the same?

3. Which of the following languages are regular, context-free but not regular, or non-context-free? Assume fixed alphabets Σ and N of >1 elements. (Here a brief explanation is sufficient.)

   (a) $L_1 = \{\, ew_1ew_2ew_3e \mid e \in \Sigma,\ w_1, w_2, w_3 \in \Sigma^*,\ |w_1| = |w_2| \,\}$.

   (b) $L_2 = \{\, ew_1ew_2ew_2^Re \mid e \in \Sigma,\ w_1, w_2, w_3 \in \Sigma^*,\ |w_1| = |w_2| \,\}$.

   (c) $L_3 = \{\, ew_1ew_2ew_3e \mid e \in \Sigma,\ w_1, w_2, w_3 \in \Sigma^*,\ |w_1| > |w_2| = |w_3| \,\}$.

   (d) $L_4 = \{\, ew_1ew_2ew_3e \mid e \in \Sigma,\ w_1, w_2, w_3 \in \Sigma^*,\ |w_1| < |w_2| \leq 6 \,\}$.

(e) $L_5 = \{\, ewewew_3e \mid e \in \Sigma,\ w, w_3 \in \Sigma^* \,\}$.

(f) $L_6 \subseteq (\Sigma \cup \{\boldsymbol{\varepsilon}, \phi, (,), +, \boldsymbol{*}\})^*$ is the set of regular expressions over Σ.

Hints: Remember that we study here regular expressions themselves, not the regular languages they define. Consider regular expressions without abbreviations (skipped parentheses).

(g) $L_7 \subseteq (\Sigma \cup N \cup \{\leftarrow\})^*$ is the set of productions of all the context-free grammars with the terminal alphabet Σ and the nonterminal alphabet N.

4. Give arguments for the context-freeness of the language

$$L_1 = \{\, a^r b^s c^r \mid r, s \in \mathbb{N} \,\}$$

and prove, using the pumping lemma, that

$$L_2 = \{\, a^r b^s c^r d^s \mid r, s \in \mathbb{N} \,\}$$

is not a context-free language.

Hint: proceed similar as with $\{a^n b^n \mid n \in \mathbb{N}\}$ and $\{a^n b^n c^n \mid n \in \mathbb{N}\}$, respectively.

5. Is the following grammar LL(1)? Is it LR(0)? LR(1)? Is it ambiguous?

$$\begin{aligned} S &\to A \\ A &\to aaAb \mid B \\ B &\to aBb \mid \epsilon \end{aligned}$$

S is the start symbol, S, A, B the nonterminal symbols, and a, b the terminal symbols of the grammar. Apply the standard method of constructing LR parsers – provide DFA's for the viable prefixes of a grammar, whose states are sets of LR(0) or LR(1) items.

As we have no general method of finding ambiguity of grammars which are not LL(1) / LR(1), in such a case an informal discussion of (un)ambiguity is sufficient.

6. Order the following formalisms (but one) according to their expressive power: placing A before B means that any language definable by A is definable by B. Also state which, if any, of them are equivalent. Point out the formalism that does not fit into the ordering.
(In this problem a brief justification – by referring to know properties – is sufficient.)

- Context-free Grammars (CFG)
- Deterministic Finite Automata (DFA)
- Deterministic Pushdown Automata (DPDA)
- LR(0) grammars
- LR(1) grammars
- Nondeterministic Finite Automata (NFA)
- Nondeterministic Finite Automata with ϵ-transitions (NFAϵ)
- Nondeterministic Turing Machines (NTM)
- Pushdown Automata (PDA)
- Regular expressions (reg.exp.)
- Turing Machines (TM)
- Turing Machines with a tape unrestricted to the right and to the left and 3 heads (TM′3)

7. (a) Show that if L is a regular language then L^R is regular.

   (b) Show that if L is a context-free language then L^R is context-free.