TDDD14/TDDD85 Formal Languages and Automata Theory
Assignment 2 – 2017
Deadline Thu. 2017-05-18, 13:00

For all the problems below it is not sufficient just to give a solution. You must **justify your answers**. In the final exam unexplained answers will be granted 0 points. You may write your solutions in english or swedish.
Hand in your solutions to
- Christer Bäckström: for TDDD85 (U1)
- Jonas Wallgren: for TDDD14 (others)

Please mark clearly which course you are taking. You have three options:
- Hand in your solutions at a lecture or tutorial.
- Put your solutions in the box "Post till IDA" in front of the Café Java in the B building.
- Email a single pdf file (no other formats, please) to the appropriate teacher and start the subject line with the course code.

1. Construct context-free grammars for the languages below. Explain which strings are generated by each variable (i.e. non-terminal symbol) of your grammar and what the role of each production is.

   (a) $L_{np} = \{x \in \{a,b\}^* \mid x \neq x^R\}$,
   where x^R is the reverse of the string x. That is, L_{np} is the complement of the language of palindromes over the alphabet $\{a,b\}$.

   (b) The language $L_3 \subseteq \{a,b,c\}^*$ of those strings of odd length where the first, middle and last symbols are the same.

2. Let $L^R = \{x^R \mid x \in L\}$. Which of the following pairs of languages differ for some choice of L or of L_1 and L_2?

$$\begin{array}{rcl} LL & \text{and} & \{xx \mid x \in L\} \\ L_1^R L_2^R & \text{and} & (L_1L_2)^R \\ L_1^R L_2^R & \text{and} & (L_2L_1)^R \\ L^* & \text{and} & L^*L \end{array}$$

   For the pairs that differ, can you find special cases where they do not differ?

3. Use the pumping lemma for context-free languages to prove that each of the following languages is not context free.

   (a) $L_1 = \{a^n b^m c^k \mid 0 \leq n < m < k\}$

   (b) Consider the alphabet $\Sigma = \{0, 1, \ldots, 9\}$ and the language L_2 defined as

   $$L_2 = \{wuv \mid w + u = v\},$$

   where the substrings w, u and v are interpreted as ordinary integers. For instance, the string $12719 \in L$ since $12 + 7 = 19$, and the string $10^n 20^n 30^n \in L$ for all $n \geq 0$ (since $1+2 = 3$, $10+20 = 30$, $100+200 = 300$ etc.).

4. Let G be the following CFG (where S is the start symbol):

   $$\begin{array}{lcl} S & \rightarrow & aB|aDc \\ B & \rightarrow & bBc|c \\ D & \rightarrow & bc|c \end{array}$$

   (a) Describe the language $L(G)$.

   (b) Show that G is ambiguous.

   (c) Show that G is not an LR(1) grammar.

   (d) Modify G into an equivalent grammar G' that is LR(1). Explain why G' is an LR(1) grammar and why G and G' are equivalent.

5. Assume some alphabet Σ and two languages $A, B \subseteq \Sigma^*$. Define the language $A \setminus B = \{x \in \Sigma^* \mid xy \in A \text{ for some } y \in B\}$. Prove that if both A and B are recursively enumerable then also $A \setminus B$ is recursively enumerable. (The term *recursively enumerable* is also called *semi-decidable* and *Turing recognizable*).

6. The mapping reduction $\leq_m$ can be viewed as a relation over languages. Prove or disprove that it has each of the following properties:

   (a) Reflexive: $L \leq_m L$ for all languages L.

   (b) Symmetric: For all languages L_1 and L_2, if $L_1 \leq_m L_2$, then $L_2 \leq_m L_1$.

   (c) Transitive: For all languages L_1, L_2 and L_3, if $L_1 \leq_m L_2$ and $L_2 \leq_m L_3$, then $L_1 \leq_m L_3$.