Homework 3

Out: 12 Feb, 2009
Due: 19 Feb., 2009

Note: Questions marked with an asterisk (*) are to be handed in. The others are for practice and will not be graded. Put your solutions to the (*) problems in the (now unique) homework box on Soda level 2 by 4pm on the due date. The usual remarks about clear answers and the collaboration policy still hold. Depending on grading resources, we may grade only a random subset of the problems and check off the rest; so you are advised to attempt all questions.

1. Which of the following languages are regular? If the language is regular, exhibit a finite automaton or a regular expression for it. If not, give a careful proof using the pumping lemma.
   (a) (*) The set of all strings over the alphabet \{ (, ) \} that consist of correctly nested pairs of parentheses. (E.g., the string ‘((())())’ belongs to this language, but the strings ‘())(‘ and ‘((’ do not.)
   (b) (*) The set of all words over the alphabet \{ a, b \} in which the number of occurrences of “abb” and of “bba” are the same. [Note: The string abba contains one occurrence of each.]
   (c) The language \( A = \{ w \in \{ 0, 1 \}^* : w = 1^k y, \text{ where } y \in \{ 0, 1 \}^* \text{ contains at least } k \text{'s, for some } k \geq 1 \} \). Note that 1101010 \( \in A \) for both \( k = 1 \) (its a 1 followed by 101010) and \( k = 2 \) (its a 11 followed by 01010), but 100 \( \notin A \) because there is no \( k \) for which the def. applies.
   (d) The language \( B = \{ w \in \{ 0, 1 \}^* : w = 1^k y, \text{ where } y \in \{ 0, 1 \}^* \text{ contains at most } k \text{'s, for some } k \geq 1 \} \).
   (e) (*) The set of all words over the alphabet \{ a, b \} in which the number of occurrences of “aaa” and of “bbb” are the same. [Note: The string bbbaaaabbb contains two occurrences of each.]
   (f) The language \{0^i1^j : i, j \geq 0 \text{ and } i \neq j \}. [HINT: Why can’t you use the pumping lemma in this case? It might be helpful to consider the complement of the language.]

2. (More closure) Let \( L \) be a regular language. Show that the following languages are regular.
   (a) (*) The language \( \text{min}(L) = \{ x \in L : \text{no proper prefix of } x \text{ is in } L \} \).
   (b) The language \( L_{10} \) consisting of the lexicographically first 10 strings of \( L \).
   (c) The language \( \frac{1}{2}(L) = \{ x : \exists y \text{ s.t. } xy \in L \text{ and } |x| = |y| \} \). [HINT: This part is quite challenging. Given a DFA for \( L \), construct a NFA for \( \frac{1}{2}(L) \). The “product construction” used in the proof that regular languages are closed under intersection should be useful here.]
   (d) (*) The language DROP-OUT\( (L) \). For a language \( L \), DROP-OUT\( (L) = \{ xz : xyz \in L, \text{ where } x, z \in \Sigma^*, y \in \Sigma \} \); i.e., its the language of all strings that can be obtained by removing one symbol from a string in \( L \).

3. Let \( M \) be a DFA with \( n \) states. Prove that the language accepted by \( M \) is infinite if and only if \( M \) accepts some string of length at least \( n \) and less than \( 2n \).
4. Which of the following statements are true? If the statement is true, provide a proof; if it is false, provide a simple counterexample.

(a) If the language $L$ contains a regular language $L'$, then $L$ is regular.
(b) (*) If $L_1$ and $L_2$ are not regular, then $L_1 \cap L_2$ is not regular.
(c) (*) If $L_1, L_2, L_3 \ldots$ are all regular, then the language $\bigcup_{i=1}^{\infty} L_i$ is also regular.