CS 381 - Spring 2008

Assignment 2: Languages, recursive definitions, and regular expressions

Total points: 100

Due: Thursday 2/21/08 at 11:30AM (beginning of class)

Complete the following exercises and submit a single set of typed or fully-legible handwritten solutions. You must provide a justification for every answer. Correct but unproved answers will cost you dearly.

1. (20 points) Let *PAL* denote the set of all palindromes over the alphabet `{a,b}`, and let PAL^k denote the subset of *PAL* containing all the palindromes of length k.
   - (5 points) Prove by construction that $|PAL^3| = |PAL^4|$.
   - (5 points) What is the value of $|PAL^{2n}|$, where n is any natural number? Prove your answer.
   - (5 points) Give a recursive/inductive definition of the set *PAL*.
   - (5 points) Give a recursive/inductive definition of the set PAL^{even}, that is, of the set of all even-length palindromes.
2. (10 points) Suppose that L is a language such that the concatenation of two words in L always yields another word in L if and only if the two words are not identical. That is, for any distinct words w_1 and w_2 in L, $w_1 w_2 \in L$ but $w_1 w_1 \notin L$. Prove that such a language L cannot exist.
3. (5 points) Let L = `{aa, ba, aba, abaab}`. Find the **shortest** string that disproves the correctness of the following algorithm for testing whether any string w of `a`'s and `b`'s is in L^*.
   Step 1: Cross off the longest prefix of w that is a word in L and set w to what is left after this prefix is removed from w.
   Step 2: Repeat step 1 until w is equal to Λ, in which case we conclude that the original $w \in L^*$, or until $w \neq \Lambda$ but no word in L is a prefix of w, in which case we conclude that the original $w \notin L^*$.
   A sample run of this algorithm on the string w_1 = `abaabaaababa`, yields the strings `aaababa`, `ababa`, `ba`, and Λ. Therefore, according to this algorithm, $w_1 \in L^*$.
4. (15 points) A language L is *closed under concatenation* if the concatenation of any two words in L (including a word concatenated with itself) is also in L.
   - (5 points) Prove that for any language L, L^* is closed under concatenation.
   - (10 points) We say that L_1 is *smaller than* L_2 if L_1 is a proper subset of L_2, denoted $L_1 \subset L_2$. Assume that $L_1 \subseteq L_2$ and L_2 is closed under concatenation. Prove that, if $L_2 \neq L_1^*$, then $L_1^* \subset L_2$. In other words, L_1^* is the smallest closed language that contains L_1.
5. (10 points) Write a recursive/inductive definition of EVEN that is different from the ones we discussed in class and those given in the book, and with which it is **trivial** to prove the statement: "All the numbers in EVEN end in the digit 0, 2, 4, 6, or 8." For full credit, you must provide the definition, prove its correctness, and write down the proof of the above statement.
6. (10 points) Write a recursive/inductive definition for each of the following two languages over the alphabet `{0,1}`.
   - (5 points) The language `ONEOONE` of binary words that contain the substring 101.
   - (5 points) The language `NOTONEOONE` of binary words that do *not* contain the substring 101.
7. (20 points) For full credit, construct the **shortest** regular expression to define each of the following languages over the alphabet `{a,b}`.
   - (5 points) All words in which each clump of `a`'s (if any) has length k, where k is a multiple of 3, e.g., `bbbb`, `aaa`, or `baaabbaaaaaab`.
   - (5 points) All words in which the *total* number of `a`'s is a multiple of 3, but where each clump of `a`'s (if any) may have any length, e.g., `bbbb` or `abababb`.
   - (5 points) All words that contain exactly two `b`'s or exactly three `b`'s, and any number of `a`'s.
   - (5 points) All words that do not contain the substring `ab`.
8. (10 points) Write a concise English description of the language defined by each of the following regular expressions. For full credit, your description must include all strings in the language and exclude all strings not in the language.
   - (5 points) $(a(a+bb)^*)^*$
   - (5 points) $(a(aa)^*b(bb)^*)^*$