CLEMSON UNIVERSITY, SCHOOL OF COMPUTING
CPSC 3500 FOUNDATIONS OF COMPUTER SCIENCE

Assignment 5: Regular Expressions

Return by 10/9/2018

Problem 1 10%

In each case below, find a string of minimum length in $\{a, b\}^*$ not in the language corresponding to the given regular expression.

(a) $b^*(ab)^*a^*$

(b) $(a^* + b^*)(a^* + b^*)(a^* + b^*)$

(c) $a^*(baa^*)^*b^*$

(d) $b^*(a + ba)^*b^*$

Problem 2 10%

Consider the two regular expressions

$$r = a^* + b^* \text{ and } s = ab^* + ba^* + b^*a + (a^*b)^*$$

(a) Find a string corresponding to r but not to s.

(b) Find a string corresponding to s but not to r.

(c) Find a string corresponding to both r and s.

(d) Find a string in $\{a, b\}^*$ corresponding to neither r nor s.

Problem 3 10%

Let r and s be arbitrary regular expressions over the alphabet Σ. In each case below, find a simpler equivalent regular expression.

(a) $r(r^*r + r^*) + r^*$

(b) $(r + \Lambda)^*$

(c) $(r + s)^*rs(r + s)^* + s^*r^*$

Problem 4 10%

It is not difficult to show using mathematical induction that for every integer $n \geq 2$, there are non-negative integers i and j such that $n = 2i + 3j$. With this in mind, simplify the regular expression $(aa + aaa)(aa + aaa)^*$.

Problem 5 10%

Suppose w and z are strings in $\{a, b\}^*$. Find regular expressions corresponding to each of the languages defined recursively below.

(a) $\Lambda \in L$; $\forall x \in L$, wx and xz are elements of L.

(b) $a \in L$; $\forall x \in L$; wx, xw, and xz are elements of L.

(c) $\Lambda \in L$; $a \in L$; $\forall x \in L$, wx and zx are in L.

Problem 6 30%

Find a regular expression corresponding to each of the following subsets of $\{a, b\}^*$.

(a) The language of all strings containing exactly two a's.

(b) The language of all strings containing at least two a's.

(c) The language of all strings that do not end with ab.

(d) The language of all strings that begin or end with aa or bb.

(e) The language of all strings not containing the substring aa.

(f) The language of all strings in which the number of a's is even.

(g) The language of all strings containing no more than one occurrence of the string aa. (The string aaa should be viewed as containing two occurrences of aa.)

(h) The language of all strings in which every a is followed immediately by bb.

(i) The language of all strings containing both bb and aba as substrings.

(j) The language of all strings not containing the substring aaa.

Problem 7 10%

(a) The regular expression $(b+ab)^*(a+ab)^*$ describes the set of all strings in $\{a,b\}^*$ not containing the substring ___x___ for any x. (Fill in the blanks appropriately.)

(b) The regular expression $(a+b)^*(aa^*bb^*aa^* + bb^*aa^*bb^*)(a+b)^*$ describes the set of all strings in $\{a,b\}^*$ containing both the substrings ___ and ___. (Fill in the blanks appropriately.)

Problem 8 10%

Prove that every finite language is regular. Hint: use induction.

Problem 9 10%

Given an FA $M = (Q, \Sigma, q_0, A, \delta)$ that accepts language $L(M)$. Describe two algorithms for construction of (not necessarily optimal) FA's that accept $\overline{L(M)}$.

Solutions to Assignment 5

Solution to Problem 1

(a) aab or abb

(b) abab or baba

(c) bba

(d) abba

Solution to Problem 2

(a) aa

(b) ba

(c) a

(d) aba

Solution to Problem 3

(a) r^*

(b) r^*

(c) $(r+s)^*$

Solution to Problem 4

Answer: aaa^*. This regular expression corresponds to the representations of all integers as sum $2i+3j$. The number of a's equals this integer.

Solution to Problem 5

(a) $w^* z^*$

(b) $w^* a(w+z)*$

(c) $(w+z)^*(a+\Lambda)$

Solution to Problem 6

(a) $b^* ab^* ab^*$

(b) Every expression of the form $AaBaC$, where each of A, B, C is either b^* or $(a+b)^*$, and at least one of the three is $(a+b)^*$, is a solution.

(c) $\Lambda + b + (a+b)^* a + (a+b)^* bb$

(d) $(aa+bb)(a+b)^* + (a+b)^*(aa+bb)$

(e) $(b+ab)^*(\Lambda+a)$ or $(\Lambda+a)(b+ba)^*$.

(f) $b^*(ab^*ab^*)^*$.

(g) The regular expression $r = (b + ab)^*$ corresponds to the set of strings that don't end with a and don't contain aa, and $s = (b + ba)^*$ to the set of strings that don't begin with a and don't contain aa. In a string with exactly one occurrence of aa, the strings before and after aa correspond to r and s, respectively. Therefore, one answer is $(b + ab)^*(a + \Lambda) + (b + ab)^* aa(b + ba)^*$. A more concise answer is $(b + ab)^*(\Lambda + a + aa)(b + ba)^*$.

(h) $(b + abb)^*$

(i) $(a + b)^*(bb(a + b)^* aba + aba(a + b)^* bb)(a + b)^*$

(j) $(\Lambda + a + aa)(b + ba + baa)^*$.

Solution to Problem 7

(a) not containing the substring $aaxbb$

(b) both the substrings ab, and ba

Solution to Problem 8
Use induction on the length of string x, and definition of regular languages.