Theory of Computation — CSE 105

Regular Languages
Study Guide and Homework I

Homework I: Solutions to the following problems should be turned in **class** on **July 12, 1999.**
Instructions:

- Write your answers clearly and completely. Please use 8.5×11 inches paper. Use a stapler or a clip to attach the individual pages. Write your name.

- When presenting any construction, for example, an algorithm or an automaton, please give an overview of the main ideas and then present the construction. Always support the correctness of your construction with a short informal proof.

1. For each of the languages given below, design finite state automata and regular expressions to recognize them. In all cases the alphabet is $\{0, 1\}$.

    (a) $L_1 = \{w | w$ does not contain the substring $110\}$
    (b) $L_2 = \{w | w$ contains an even number of 0's, or exactly two 1's $\}$

2. Problem 1.26, Page 88.

3. Problem 1.42, Page 90. Provide a short proof of the correctness of your construction.

4. Prove the following languages non-regular:

    (a) $PRIMES = \{a^p | p$ is a prime$\}$;
    (b) Let

    $$\Sigma_2 = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

    Here, Σ_2 contains all columns of 0's and 1's of height two. A string of symbols in Σ_2 gives two rows of 0's and 1's. Consider each row to be a binary number and let

    $$E = \{w \in \Sigma_2 | \text{the bottom row of } w \text{ is the reverse of the top row of } w\}.$$

Study Guide: In the following, the material on regular languages (chapter 1) is broken down into a number of short topics. For each topic, a list of specific items and problems are provided. If you understand these items and solve the problems, you would do very well in the course.

1 Deterministic Finite Automata (DFA)

Topics: The notion and definition of DFA, presentation of DFA by transition diagrams, the notion of acceptance by a DFA, the class of regular languages, techniques for designing DFAs, and closure operations.

Designing DFAs

1. 1.1, 1.2, and 1.3, pages 83 and 84.

2. Exercise 1.4, Page 84.

3. For each of the following regular expressions, draw a DFA recognizing the corresponding language.

   (a) $(0 \cup 1)^*110^*$
   (b) $(11 \cup 10)^*$
   (c) $(1 \cup 110)^*0$

4. Draw a DFA that recognizes the language of all strings of 0's and 1's of length ≥ 1 that, if they were interpreted as binary representations of integers, would represent integers evenly divisible by 3. Leading 0's are permissible.

5. Show that if L is a regular language and F is finite language, then $L \cup F$, $L \cap F$, and $L - F$ are regular.

6. Show that if L is a non–regular language and F is a finite language, then $L \cup F$ and $L - F$ are non–regular.

7. Problems 1.25 and 1.27, Page 88.

8. 1.29, 1.30, 1.41

Closure Properties of Regular Languages

1. For each statement below, decide whether it is true or false, If it is true, prove it; if not, give a counter example. All parts refer to languages over $\{a, b\}$.

   (a) If $L_1 \subseteq L_2$ and L_1 is not regular, then L_2 is not regular.
   (b) If $L_1 \subseteq L_2$ is not regular, then L_1 is regular.
   (c) If L_1 and L_2 are nonregular, then $L_1 \cup L_2$ is nonregular.

(d) If L_1 and L_2 are nonregular, then $L_1 \cap L_2$ is nonregular.

(e) If L is not regular, then $\bar{L}$, the complement of L, is not regular.

(f) If L_1 is regular and L_2 is nonregular, then $L_1 \cup L_2$ is nonregular.

(g) If L_1 is regular, L_2 is nonregular, and $L_1 \cap L_2$ is nonregular, then $L_1 \cup L_2$ is nonregular.

(h) If $L_1, L_2, \cdots$ are regular, then $\cup_{n=1}^{\infty} L_n$ is regular.

2. Problem 1.24, 1.42

2   Nondeterministic Finite Automata (NFA)

Topics: The notion of nondeterminism, definition of acceptance for NFAs, economy of states by using NFA, equivalence of DFAs and NFAs, and examples that illustrate the conversion of an NFA to an equivalent DFA.

Notion of Nondeterminism 1.9, 1.10, page 85

Practice in Designing NFAs 1.5, 1.6, 1.7, 1.8, pages 84 and 85.

Practice in converting an NFA to an equivalent DFA 1.12, page 85.

3   Regular Expressions (RE)

Topics: The definition of regular expressions, writing regular expressions, equivalence with finite automata: every regular expression has an equivalent finite automata and every finite automata has an equivalent regular expression.

Basics of Regular Expressions

1. What is the shortest string of a's and b's not in the language corresponding to the regular expression $b^*(abb^*)^*a^*$?

2. Consider the following two regular expressions $R_1 = a^* + b^*$ and $R_2 = ab^* + ba^* + b^*a + (a^*b)^*$.

   (a) Find a string corresponding to R_1 but not to R_2.

   (b) Find a string corresponding to R_2 but not to R_1.

   (c) Find a string corresponding to both R_1 and R_2.

   (d) Find a string corresponding to neither R_1 nor R_2.

3. Simplify the following regular expressions:

(a) $(01 \cup 10 \cup 0110 \cup 1001)^*$

(b) $(0(0 \cup 1)^*)^+$

(c) $01((01)^*01 \cup (01)^*) \cup (01)^*$

(d) $(01 \cup \varepsilon)^*$

(e) $(10 \cup 01)^*1001(10 \cup 01)^* \cup (01)^*(10)^*$

4. What is true of the language corresponding to a regular expression that does not involve the operators $*$ or $+$? Why?

Designing Regular Expressions

1. 1.13, page 86

2. Find regular expressions corresponding to each of the languages defined recursively below.

   (a) $\varepsilon \in L$; if $x \in L$, then $aabx$ and xbb are elements of L; nothing is in L unless it can be obtained from these two statements.

   (b) $a \in L$; if $x \in L$, then $aabx$, $xaab$, and xbb are elements of L; nothing is in L unless it can be obtained from these two statements.

3. Find a regular expression corresponding to each of the following subsets of $\{0, 1\}^*$.

   (a) The language of strings containing exactly two 0's.

   (b) The language of strings containing at least two 0's.

   (c) The language of strings that do not end with 01.

   (d) The language of strings that begin or end with 00 or 11.

   (e) The language of strings containing no more than one occurrence of the string 00. (The string 000 should be viewed as containing two occurrences of 00.)

   (f) The language of strings in which the number of 0's is even.

   (g) The language of strings in which every 0 is immediately followed by 11.

   (h) The language of strings that do not contain the substring 110.

   (i) The language of strings that do contain both the substring 11 and the substring 010.

Interpreting Regular Expressions Describe as simply as possible the language corresponding to each of the following regular expressions.

1. $0^*1(0^*10^*1)^*0^*$

2. $((0 \cup 1)^3)^*(\varepsilon \cup 0 \cup 1)$

3. $(1 \cup 01)^*(0 \cup 01)^*$

4. $(0 \cup 1)^*(0^+1^+0^+ \cup 1^+0^+1^+)(0 \cup 1)^*$

Practice the Translation Algorithm from REs to NFAs 1.14, page 86

Practice the Translation Algorithm from DFAs to REs 1.16, page 86

4   Non–regular Languages

Topics: Pumping lemma, examples of nonregular languages and applications of pumping lemma.

Application of Pumping Lemma

1. Using the Pumping Lemma show that each of these languages is not regular.

   (a) $L = \{a^n b a^{2n} | n \geq 0\}$

   (b) $L = \{a^i b^j c^k | k > i + j\}$

   (c) $L = \{x \in \{a, b\}^* | N_a(x) < 2N_b(x)\}$ where $N_a(x)$ ($N_b(x)$) is the number of occurrences of the letter a (b) in x.

   (d) $L = \{x \in \{a, b\}^* | $ no initial substring of x has more b's than a's$\}$

   (e) $L = \{x \in \{a, b\}^* | x$ is a palindrome$\}$

   (f) $L = \{ww | w \in \{a, b\}^*\}$

2. Here is a 'proof' using the pumping lemma, that the language L of all strings of a's and b's of length 100 is not regular. Since the result being 'proved' is false (all finite languages are regular), the proof cannot be correct. What is the flaw in the proof?

   Assume that L is regular. By the pumping lemma, if we choose an element of L, say $w = a^1 00$, there are string x, y, and z, with $|y| > 0$, so that every string of the form $xy^k z$ (where $k \geq 0$) is in L. Since there are infinitely many different strings of this form, this contradicts the fact that L is finite. Therefore, L is not regular.

3. 1.17, page 86.

4. 1.23, 1.28, 1.33, 1.36, pages 88 and 89

5. 1.38, page 90

6. 1.40, 1.43, page 90

Regular or Nonregular Below are a number of languages over $\{a, b\}$. In each case, decide whether the language is regular or not, and prove that your answer is correct.

1. L is the set of strings X beginning with a non-null string of the form ww.

2. L is the set of all strings x having some non–null string of the form ww.

3. L is the set of strings x having some non–null substring of the form www.

4. $L = \{x \in \{a,b\}^* | x$ is not a palindrome$\}$

5. $L = \{x \in \{a,b\}^* | x$ begins with a palindrome of length $\geq 3 \}$

6. $L = \{x \in \{a,b\}^* | \ N_a(x)$ is a perfect square$\}$

7. $L = \{x \in \{a,b\}^* | $ in every initial string of $x,$ the number of a's and the number of b's differ by no more than 2$\}$

8. $L = \{x \in \{a,b\}^* | $ in every substring of x, the number of a's and the number of b's differ by no more than 2$\}$

9. $L = \{x \in \{a,b\}^* | N_a(x)$ and $N_b(x)$ are both divisible by 5$\}$

10. $L = \{x \in \{a,b\}^* | $ there is some integer $k > 1$ so that $N_a(x)$ and $N_b(x)$ are both divisible by $k \}$

5   Decision Algorithms

Describe decision algorithms to answer each of the following questions.

1. Given two DFAs M_1 and M_2, are there any strings that are accepted by neither?

2. Given an NFA M and a string x, does M accept x?

3. Given two NFAs, do they accept the same language?

4. Given an NFA M and a string x, is there more than one sequence of transitions corresponding to x that causes M to accept x?

6   Miscellaneous Problems

1. 1.31

2. **Myhill-Nerode Theorem:** 1.34 and 1.35

3. **Number of states:** 1.39 and 1.44

4. **Transducers:** 1.19, 120, 1.21 and 1.22