CS Theory (Spring '24) — Assigned: February 5, 2024

Homework 2

Instructor: *Josh Alman* — Due: February 14, 2024 at 11:59 pm

Collaboration on homework is allowed. However, you must write up solutions by yourself and understand everything that you hand in. You may also consult other reference materials, but you may not seek out answers from other sources.

Write the solution to each problem on a separate page. Please do not write your name on any page of the submission; we are using anonymous grading in GradeScope.

Be sure to answer the final question listing your collaborators and other sources used. See the course webpage for more details.

There are 120 total possible points, and also bonus questions worth an additional 20 points.

1 Infinite Things (20 points; each part 10 points)

(a) On post #32 in Ed Discussion, our TA William wrote that "regular languages aren't closed under INFINITE unions. Otherwise every language would be regular (which isn't true)." Prove William's claim. In other words, prove that for any language L, we can find an infinite sequence of regular languages $L_1, L_2, L_3, \ldots$ such that $L = L_1 \cup L_2 \cup L_3 \cup \cdots$.

(b) Recall that in a DFA, we require the set of states to be a finite set. Let's define a DIA (Deterministic Infinite Automaton) in the same way as a DFA, except that the set of states may be any set (not necessarily finite). Prove that every language is recognized by a DIA. Be sure to completely and unambiguously describe your DIA, for instance as a 5-tuple.

(c) (BONUS, 10 additional points) Suppose L is an infinite regular language. Prove that there are two infinite regular languages L_1, L_2 with $L_1 \cup L_2 = L$ and $L_1 \cap L_2 = \emptyset$.

2 Is it regular? (30 points; each part 10 points)

For each of the following languages over $\Sigma = \{0, 1\}$, determine (with proof) whether it is regular. (At least one is regular, and at least one isn't!)

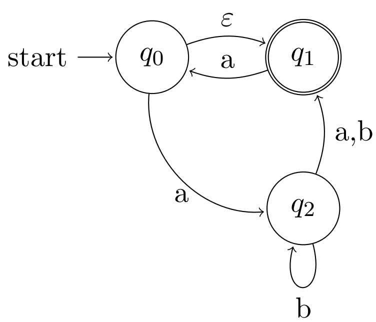
(a) $\{0^n 1^m \mid n + m \geq 4\}$

(b) $\{0^n 1^m \mid n - m \geq 4\}$

(c) $\{0^n 1^m \mid m \text{ is a prime number, and } n \geq 4\}$

3 3-state NFAs (30 points; each part 10 points)

Let L_1 be the language recognized by the following NFA over the alphabet $\Sigma = \{a,b\}$:



(a) Give a DFA that recognizes L_1, and explain why it's correct.

(b) Give a regular expression for L_1, and explain why it's correct.

(c) Consider the language $L_2 := \{a^x b^y c^z \mid x \geq 0, y > 0, z \geq 0\}$ over the alphabet $\Sigma = \{a, b, c\}$. Give a 3-state NFA for L_2 and explain why it's correct.

4 Closure Properties?? (30 points; each part 10 points)

Prove or disprove each of the following claims. (At least one of the first three is true and at least one of the first three is false!)

(a) If L_1 and L_2 are regular languages over alphabet Σ, then

$$\{w \in \Sigma^* \mid w \text{ is in exactly one of } L_1 \text{ and } L_2\}$$

is a regular language. ("Exactly one" means it can't be in both and can't be in neither.)

(b) If L is a language over alphabet Σ and L^* is a regular language, then L is a regular language.

(c) Say two strings $x, y \in \Sigma^*$ over alphabet Σ are *typos* if $|x| = |y|$ and they differ in exactly one position. For example, 00000 and 01000 are typos (they differ in the second position), but 111 and 111 are not typos (they don't differ anywhere), 111111 and 110011 are not typos (they differ in two positions), and 101 and 10 are not typos (they have different lengths). If L is a regular language over alphabet Σ, then

$$\{y \in \Sigma^* \mid \text{ there is a } x \in L \text{ such that } x, y \text{ are typos}\}$$

is a regular language.

(d) (BONUS, 10 additional points) Recall from the first homework that if w is a string, then w^R denotes its reverse. If L is a regular language over alphabet Σ, then

$$\{w \in \Sigma^* \mid ww^R \in L\}$$

is a regular language.

5 Day-to-Day Regular Expression (9 points)

Many tools used by software engineers (including command-line utilities like grep, and text editors like Emacs and Vim) allow one to search for substrings of a given file that match a given regular expression. Give an example of a reasonable situation where a software engineer might want to use this, and a "find" (Ctrl+f) would not suffice. (For example, your regular expression may help the engineer to write, read, or debug code.)

6 Collaborators and References (1 point)

Please list who you collaborated with on each problem, including any TAs or students you discussed the problems with in office hours. Also list any reference materials consulted other than the lectures and textbook for our class.