GROUP HOMEWORK 6, CPSC 421/501, FALL 2025

JOEL FRIEDMAN

Disclaimer: The material may sketchy and/or contain errors, which I will elaborate upon and/or correct in class. For those not in CPSC 421/501: use this material at your own risk...

Please note:

(1) You must justify all answers; no credit is given for a correct answer without justification.
(2) Proofs should be written out formally.
(3) You do not have to use LaTeX for homework, but **homework that is too difficult to read will not be graded.**
(4) You may work together on homework in groups of up to four, **but you must submit a single homework as a group submission under Gradescope.**

(0) Who are your group members? Please print if writing by hand.

(1) Let $\Sigma = \{a, b\}$.
   (a) Let L be the language of words over $\Sigma = \{a, b\}$ that end in an a; hence each string in L has length at least one, and the first few strings of L (in order of length, breaking ties with lexicographical order) are
$$L = \{a, aa, ba, aaa, aba, \ldots\}.$$
   Construct a two-state DFA that recognizes L, and explain why your DFA recognizes L. [Hint: Your DFA will have to "remember" the last symbol in the string that has been read up to that point. Hence it seems that[1] the DFA needs at least two states. Are two states sufficient?]
   (b) Let L be the language of words over $\Sigma = \{a, b\}$ whose second to last letter is a; hence each string in L has length at least two, and the first few strings of L (in order of length, breaking ties with lexicographical order) are
$$L = \{aa, ab, aaa, aab, baa, bab, aaaa, \ldots\}.$$

[1] Once we cover the Myhill-Nerode theorem, we will be able to make this argument rigorous.

Construct a four-state DFA that recognizes L. [Hint: Your DFA will have to "remember" the last two symbols in the string that has been read up to that point. Hence it seems that the DFA needs at least four states. Are four states sufficient?]

(2) For each infinite increasing sequence $n_1 < n_2 < n_3 < \cdots$ of non-negative numbers, let

$$L = L_{n_1,n_2,\ldots} = \{a^{n_1}, a^{n_2}, \ldots\} \subset \{a\}^*.$$

Show that if L is regular, then for some $p \in \mathbb{N}$ we have for all $i \in \mathbb{N}$, $n_{i+1} \le n_i + p$. (We've already explained roughly why this is true in class this year, but we didn't give a formal argument.) You might refer to Figure 2 of the handout "Non-regular languages and the Myhill-Nerode Theorem,..." or a relevant diagram from our class notes.

(3) Show that there are only a countably infinite number of regular languages over a fixed alphabet, Σ. [Hint: For any $m \in \mathbb{N}$, show that there are only finitely many languages that can be recognized by DFA's with m states.]

(4) (a) Show that there are uncountably many different sequences $n_1 < n_2 < n_3 < \cdots$ such that for all $i \in \mathbb{N}$, $n_{i+1} \le n_i + 2$. [Hint: We know that Power($\mathbb{N}$) is uncountable, so it suffices to give an injection from Power($\mathbb{N}$) to the set of such sequences.]

(b) Show that there are is a sequence of non-negative integers $n_1 < n_2 < n_3 < \cdots$ such that for all $i \in \mathbb{N}$, $n_{i+1} \le n_i + 2$, and that

$$L = L_{n_1,n_2,\ldots} = \{a^{n_1}, a^{n_2}, \ldots\}$$

is not regular.

Department of Computer Science, University of British Columbia, Vancouver, BC V6T 1Z4, CANADA.

Email address: `jf@cs.ubc.ca`

URL: `http://www.cs.ubc.ca/~jf`