Fall 2021 CIS 511

Introduction to the Theory of Computation
Jean Gallier

Homework 2

September 27, 2021; Due October 11, 2021

Problem B1 (40 pts). Let $\Sigma = \{a_1, \ldots, a_n\}$ be an alphabet of n symbols, with $n \geq 2$.

(1) Construct an NFA with $2n+1$ states accepting the set L_n of strings over Σ such that, every string in L_n has an odd number of a_i, for some $a_i \in \Sigma$. Equivalently, if L_n^i is the set of all strings over Σ with an odd number of a_i, then $L_n = L_n^1 \cup \cdots \cup L_n^n$.

(2) Prove that there is a DFA with 2^n states accepting the language L_n.

(3) Prove that every DFA accepting L_n has at least 2^n states.

Hint. If a DFA D with $k < 2^n$ states accepts L_n, show that there are two strings u, v with the property that, for some $a_i \in \Sigma$, u contains an odd number of a_i's, v contains an even number of a_i's, and D ends in the same state after processing u and v. From this, conclude that D accepts incorrect strings.

Problem B2 (50 pts). Let R be any regular language over some alphabet Σ. Prove that the language

$$L = \{u \in \Sigma^* \mid \exists v \in \Sigma^*,\, uv \in R,\, |u| = |v|\}$$

is regular.
Hint. Think nondeterministically; use a (nonstandard) cross-product construction.

Problem B3 (50 pts). (a) Let $T = \{0, 1, 2\}$, let C be the set of 20 strings of length three over the alphabet T,

$$C = \{u \in T^3 \mid u \notin \{110, 111, 112, 101, 121, 011, 211\}\},$$

let $\Sigma = \{0, 1, 2, c\}$, and consider the language

$$L_M = \{w \in \Sigma^* \mid w = u_1 c u_2 c \cdots c u_n,\, n \geq 1, u_i \in C\}.$$

Prove that L_M is regular (there is a DFA with 7 states).

(b) The language L_M has a geometric interpretation as a certain subset of $\mathbb{R}^3$ (actually, $\mathbb{Q}^3$), as follows: Given any string, $w = u_1 c u_2 c \cdots c u_n \in L_M$, denoting the jth character in u_i

by u_i^j, where $j \in \{1, 2, 3\}$, we obtain three strings

$$\begin{aligned} w^1 &= u_1^1 u_2^1 \cdots u_n^1 \\ w^2 &= u_1^2 u_2^2 \cdots u_n^2 \\ w^3 &= u_1^3 u_2^3 \cdots u_n^3. \end{aligned}$$

For example, if $w = 012c001c222c122$ we have $w^1 = 0021$, $w^2 = 1022$, and $w^3 = 2122$. Now, a string $v \in T^+$ can be interpreted as a decimal real number written in base three! Indeed, if

$$v = b_1 b_2 \cdots b_k, \quad \text{where} \quad b_i \in \{0, 1, 2\} = T \ (1 \leq i \leq k),$$

we interpret v as $n(v) = 0.b_1 b_2 \cdots b_k$, i.e.,

$$n(v) = b_1 3^{-1} + b_2 3^{-2} + \cdots + b_k 3^{-k}.$$

Finally, a string, $w = u_1 c u_2 c \cdots c u_n \in L_M$, is interpreted as the point, $(x_w, y_w, z_w) \in \mathbb{R}^3$, where

$$x_w = n(w^1), \ y_w = n(w^2), \ z_w = n(w^3).$$

Therefore, the language, L_M, is the encoding of a set of rational points in $\mathbb{R}^3$, call it M. This turns out to be the part consisting of the rational points having a finite decimal representation in base 3 of a fractal known as the *Menger sponge*.

Describe recursive rules to create the set M, starting from a unit cube in $\mathbb{R}^3$. Justify as best as you can how these rules are derived from the description of the coordinates of the points of M defined above (which points are omitted, included, ...).

Draw some pictures illustrating this process and showing approximations of the Menger sponge.

Extra Credit (30 points). Write a computer program to draw the Menger sponge (based on the ideas above).

Problem B4 (60 pts). Recall from class that given any DFA $D = (Q, \Sigma, \delta, q_0, F)$, a *congruence* $\equiv$ *on* D is an equivalence relation $\equiv$ on Q satisfying the following conditions:

(1) For all $p, q \in Q$ and all $a \in \Sigma$, if $p \equiv q$, then $\delta(p, a) \equiv \delta(q, a)$.

(2) For all $p, q \in Q$, if $p \equiv q$ and $p \in F$, then $q \in F$.

(a) Given a congruence $\equiv$ on a DFA D, we define the *quotient DFA* $D/\equiv$ as follows: denoting the equivalence class of a state $p \in Q$ as $[p]$,

$$D/\equiv \, = (Q/\equiv, \Sigma, \delta/\equiv, [q_0], F/\equiv),$$

where the transition function $\delta/\equiv$ is given by

$$\delta/\equiv ([p], a) = [\delta(p, a)]$$

for all $p \in Q$ and all $a \in \Sigma$.

Why is $D/\equiv$ well defined? Prove that there is a surjective proper homomorphism $\pi\colon D \to D/\equiv$, and thus, that $L(D) = L(D/\equiv)$ (you may use results from HW1).

(b) Given a DFA D, prove that the state equivalence relation $\equiv_D$ is the coarsest congruence on D (this means that if $\equiv$ is any congruence on D, then $\equiv \subseteq \equiv_D$).

(c) Given two DFA's $D_1 = (Q_1, \Sigma, \delta_1, q_{0,1}, F_1)$ and $D_2 = (Q_2, \Sigma, \delta_2, q_{0,2}, F_2)$, with D_1 trim, prove that the following properties hold.

(1) There is a DFA morphism $h\colon D_1 \to D_2$ iff

$$\simeq_{D_1} \subseteq \simeq_{D_2} .$$

(2) There is a DFA F-map $h\colon D_1 \to D_2$ iff

$$\simeq_{D_1} \subseteq \simeq_{D_2} \quad \text{and} \quad L(D_1) \subseteq L(D_2);$$

(3) There is a DFA B-map $h\colon D_1 \to D_2$ iff

$$\simeq_{D_1} \subseteq \simeq_{D_2} \quad \text{and} \quad L(D_2) \subseteq L(D_1).$$

Furthermore, in all three cases, h is surjective iff D_2 is trim. Conclude that if D_1, D_2 are trim and $L(D_1) = L(D_2)$, then there is a unique surjective proper homomorphism $h\colon D_1 \to D_2$ iff

$$\simeq_{D_1} \subseteq \simeq_{D_2} .$$

(you may use results from HW1).

Prove that for any trim DFA D, there is a unique surjective proper homomorphism from D to any minimal DFA D_m accepting $L = L(D)$.

(d) Given a regular language L, prove that a minimal DFA D_m for L is characterized by the property that there is unique surjective proper homomorphism $h\colon D \to D_m$ from any trim DFA D accepting L to D_m.

Problem B5 (50 pts). (*Ultimate periodicity*) A subset U of the set $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ of natural numbers is *ultimately periodic* if there exist $m, p \in \mathbb{N}$, with $p \geq 1$, so that $n \in U$ iff $n + p \in U$, for all $n \geq m$.

(i) Prove that $U \subseteq \mathbb{N}$ is ultimately periodic iff either U is finite or there is a finite subset $F \subseteq \mathbb{N}$ and there are $k \leq p$ numbers $m_1, \ldots, m_k$, with $m_1 < m_2 < \cdots < m_k < m_1 + p$, and with m_1 the smallest element of U so that for some $p \geq 1$, $n \in U$ iff $n + p \in U$, for all $n \geq m_1$, so that

$$U = F \cup \bigcup_{i=1}^{k} \{m_i + jp \mid j \in \mathbb{N}\}.$$

Give an example of an ultimately periodic set U such that m and p are not necessarily unique, i.e., U is ultimately periodic with respect to m_1, p_1 and m_2, p_2, with $m_1 \neq m_2$ and $p_1 \neq p_2$.

Remark: A subset of $\mathbb{N}$ of the form $\{m + ip \mid i \in \mathbb{N}\}$ (allowing $p = 0$) is called a *linear set*, and a finite union of linear sets is called a *semilinear set*. Thus, (i) says that a set is ultimately periodic iff it is semilinear.

(ii) Let $L \subseteq \{a\}^*$ be a language over the one-letter alphabet $\{a\}$. Prove that L is a regular language iff the set $\{m \in \mathbb{N} \mid a^m \in L\}$ is ultimately periodic. Prove that the family of semilinear sets is closed under union, intersection and complementation (i.e., it is a boolean algebra).

(iii) Let $L \subseteq \Sigma^*$ be a regular language over any alphabet Σ (not necessarily consisting of a single letter). Prove that the set

$$|L| = \{|w| \mid w \in L\}$$

is ultimately periodic.

TOTAL: 250 points + 30 extra credit.