Spring, 2010 CIS 511

Introduction to the Theory of Computation
Jean Gallier

Homework 2

February 4, 2010; Due February 18, 2010

"A problems" are for practice only, and should not be turned in.

Problem A1. Recall that two regular expressions R and S are equivalent, denoted as $R \cong S$, iff they denote the same regular language $\mathcal{L}[R] = \mathcal{L}[S]$. Show that the following identities hold for regular expressions:

$$\begin{aligned} R^{**} &\cong R^* \\ (R+S)^* &\cong (R^* + S^*)^* \\ (R+S)^* &\cong (R^* S^*)^* \\ (R+S)^* &\cong (R^* S)^* R^* \end{aligned}$$

Problem A2. Recall that a homomorphism $h \colon \Sigma^* \to \Delta^*$ is a function such that $h(uv) = h(u)h(v)$ for all $u, v \in \Sigma^*$. Given any language, $L \subseteq \Sigma^*$, we define $h(L)$ as

$$h(L) = \{h(w) \mid w \in L\}.$$

Prove that if $L \subseteq \Sigma^*$ is a regular language, then so is $h(L)$.

Problem A3. Construct an NFA accepting the language $L = \{aa, aaa\}^*$. Apply the subset construction to get a DFA accepting L.

"B problems" must be turned in.

Problem B1 (25 pts). Let $\Sigma = \{a_1, \ldots, a_n\}$ be an alphabet of n symbols.

(a) Construct an NFA with $2n+1$ (or $2n$) 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 $L_n = L_n^1 \cup \cdots \cup L_n^n$.

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

(c) 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 (25 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 is regular.

(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 rational part of a fractal known as the *Menger sponge*.

Explain the best you can what are the recursive rules to create the Menger sponge, starting from a unit cube in $\mathbb{R}^3$. Draw some pictures illustrating this process and showing approximations of the Menger sponge.

Extra Credit (20 points). Write a computer program to draw the Menger sponge (based on the ideas above).

Problem B3 (50 pts). Given any two DFA's $D_1 = (Q_1, \Sigma, \delta_1, q_{0,1}, F_1)$ and $D_2 = (Q_2, \Sigma, \delta_2, q_{0,2}, F_2)$, a *morphism* $h\colon D_1 \to D_2$ *of DFA's* is a function $h\colon Q_1 \to Q_2$ satisfying the following two conditions:

(1) $h(\delta_1(p, a)) = \delta_2(h(p), a)$, for all $p \in Q_1$ and all $a \in \Sigma$;

(2) $h(q_{0,1}) = q_{0,2}$.

An *F-map $h\colon D_1 \to D_2$ of DFA's* is a morphism satisfying the condition

(3a) $h(F_1) \subseteq F_2$.

A *B-map $h\colon D_1 \to D_2$ of DFA's* is a morphism satisfying the condition

(3b) $h^{-1}(F_2) \subseteq F_1$.

A *proper homomorphism of DFA's* is an F-map of DFA's which is also a B-map of DFA's, i.e. it satisfies the condition

(3c) $h^{-1}(F_2) = F_1$.

We say that a morphism (resp. F-map, resp. B-map) $h\colon D_1 \to D_2$ is *surjective* if $h(Q_1) = Q_2$.

(a) Prove that if $f\colon D_1 \to D_2$ and $g\colon D_2 \to D_3$ are morphisms (resp. F-maps, resp B-maps) of DFAs, then $g \circ f\colon D_1 \to D_3$ is also a morphism (resp. F-map, resp B-map) of DFAs.

Prove that if $f\colon D_1 \to D_2$ is an F-map that is an isomorphism then it is also a B-map, and that if $f\colon D_1 \to D_2$ is a B-map that is an isomorphism then it is also an F-map.

(b) If $h\colon D_1 \to D_2$ is a morphism of DFA's, prove that

$$h(\delta_1^*(p, w)) = \delta_2^*(h(p), w),$$

for all $p \in Q_1$ and all $w \in \Sigma^*$.

As a consequence, prove the following facts:

If $h\colon D_1 \to D_2$ is an F-map of DFA's, then $L(D_1) \subseteq L(D_2)$. If $h\colon D_1 \to D_2$ is a B-map of DFA's, then $L(D_2) \subseteq L(D_1)$. Finally, if $h\colon D_1 \to D_2$ is a proper homomorphism of DFA's, then $L(D_1) = L(D_2)$.

(c) Let D_1 and D_2 be DFA's and assume that there is a morphism $h\colon D_1 \to D_2$. Prove that h induces a unique surjective morphism $h_r\colon (D_1)_r \to (D_2)_r$ (where $(D_1)_r$ and $(D_2)_r$ are the trim DFA's defined in problem B1). This means that if $h\colon D_1 \to D_2$ and $h'\colon D_1 \to D_2$ are DFA morphisms, then $h(p) = h'(p)$ for all $p \in (Q_1)_r$, and the restriction of h to $(D_1)_r$ is surjective onto $(D_2)_r$. Moreover, if $L(D_1) = L(D_2)$, prove that h induces a unique surjective proper homomorphism $h_r\colon (D_1)_r \to (D_2)_r$.

(d) Relax the condition that a DFA morphism $h\colon D_1 \to D_2$ maps $q_{0,1}$ to $q_{0,2}$ (so, it is possible that $h(q_{0,1}) \neq q_{0,2}$), and call such a function a *weak morphism*. We have an obvious notion of *weak F-map*, *weak B-map* and *weak proper homomorphism* (by imposing condition (3a) or condition (3b), or (3c)). For any language, $L \subseteq \Sigma^*$ and any fixed string, $u \in \Sigma^*$, let $D_u(L)$, also denoted L/u (called the *(left) derivative of L by u*), be the language

$$D_u(L) = \{v \in \Sigma^* \mid uv \in L\}.$$

Prove the following facts, **assuming that D_2 is trim**: If $h\colon D_1 \to D_2$ is a weak F-map of DFA's, then $L(D_1) \subseteq D_u(L(D_2))$, for some suitable $u \in \Sigma^*$. If $h\colon D_1 \to D_2$ is a weak B-map of DFA's, then $D_u(L(D_2)) \subseteq L(D_1)$, for the same u as above. Finally, if $h\colon D_1 \to D_2$ is a weak proper homomorphism of DFA's, then $L(D_1) = D_u(L(D_2))$, for the same u as above.

Suppose there is a weak morphism $h\colon D_1 \to D_2$. What can you say about the restriction of h to $(D_1)_r$? What can you say about surjectivity ? (you may need to consider $(D_2)_r$ with respect to a **different** start state). What happens (and what can you say) if D_2 is **not** trim?

Problem B4 (20 pts). Let R be any regular language over some alphabet Σ. Prove that the language

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

is regular

Problem B5 (50 pts). An *a-transducer* (or *nondeterministic sequential transducer with accepting states*) is a sextuple $M = (K, \Sigma, \Delta, \lambda, q_0, F)$, where K is a finite set of states, Σ is a finite input alphabet, Δ is a finite output alphabet, $q_0 \in K$ is the start (or initial) state, $F \subseteq K$ is the set of accepting (of final) states, and

$$\lambda \subseteq K \times \Sigma^* \times \Delta^* \times K$$

is a finite set of quadruples called the *transition function* of M.

An a-transducer defines a binary relation between Σ^* and Δ^*, or equivalently, a function $M\colon \Sigma^* \to 2^{\Delta^*}$. We can explain what this function is by describing how an a-transducer makes a sequence of moves from configurations to configurations. The current configuration of an a-transducer is described by a triple $(p, u, v) \in K \times \Sigma^* \times \Delta^*$, where p is the current state, u is the remaining input, and v is some ouput produced so far. We define the binary relation $\vdash_M$ on $K \times \Sigma^* \times \Delta^*$ as follows: For all $p, q \in K$, $u, \alpha \in \Sigma^*$, $\beta, v \in \Delta^*$, if $(p, u, v, q) \in \lambda$, then

$$(p,\, u\alpha,\, \beta) \vdash_M (q,\, \alpha,\, \beta v).$$

Let $\vdash_M^*$ be the transitive and reflexive closure of $\vdash_M$.

The function $M\colon \Sigma^* \to 2^{\Delta^*}$ is defined such that for every $w \in \Sigma^*$,

$$M(w) = \{y \in \Delta^* \mid (q_0,\, w,\, \epsilon) \vdash_M^* (f,\, \epsilon,\, y),\ f \in F\}.$$

For every language $L \subseteq \Sigma^*$, let

$$M(L) = \bigcup_{w \in L} M(w).$$

(a) Let $\Sigma = \Delta = \{a, b\}$. Construct an a-transducer swapping a's and b's (for instance, if $w = abbaa$, then $y = baabb$).

(b) Given an a-transducer $M = (K, \Sigma, \Delta, \lambda, q_0, F)$, define the new alphabet T as follows:

$$T = \{[p, u, v, q] \mid (p, u, v, q) \in \lambda\}.$$

Let $f: T^* \to \Sigma^*$ and $g: T^* \to \Delta^*$ be the homomorphisms defined such that

$$f([p, u, v, q]) = u, \quad \text{and} \quad g([p, u, v, q]) = v.$$

Prove that the language

$$R = \{[q_0, u_1, v_1, q_1][q_1, u_2, v_2, q_2] \cdots [q_{n-2}, u_{n-1}, v_{n-1}, q_{n-1}][q_{n-1}, u_n, v_n, q_n] \\ \mid [q_{i-1}, u_i, v_i, q_i] \in T,\ 1 \leq i \leq n,\ q_n \in F,\ n \geq 1\} \cup \{\epsilon \mid q_0 \in F\}$$

is a regular language.

(c) Prove that

$$f^{-1}(L) \cap R = \{[q_0, u_1, v_1, q_1][q_1, u_2, v_2, q_2] \cdots [q_{n-2}, u_{n-1}, v_{n-1}, q_{n-1}][q_{n-1}, u_n, v_n, q_n] \\ \mid [q_{i-1}, u_i, v_i, q_i] \in T,\ u_1 u_2 \cdots u_n \in L,\ q_n \in F,\ n \geq 1\} \cup \{\epsilon \mid q_0 \in F,\ \epsilon \in L\}.$$

(d) Prove that

$$M(L) = g(f^{-1}(L) \cap R).$$

If $\mathcal{L}$ is a family of languages closed under intersection with regular languages, homomorphic images, and inverse homomorphic images, is $\mathcal{L}$ closed under a-transductions? (Justify your answer).

If L is a regular language, is $M(L)$ regular? (Justify your answer).

(e) If M is an a-transducer from Σ^* to Δ^* prove that for any regular language, $L' \subseteq \Delta^*$, the language $M^{-1}(L')$ is also regular (see the definition of $M^{-1}(L')$ in the class notes).

TOTAL: 170 + 20 points.