

Introduction to the Theory of Computation

Homework 3

February 20, 2003; Due March 18, beginning of class

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

Problem A1. Prove that every finite language is regular.

Problem A2. Sketch an algorithm for deciding whether two regular expressions R, S are equivalent (i.e, whether $\mathcal{L}[R] = \mathcal{L}[S]$).

Problem A3. Given any language $L \subseteq \Sigma^*$, let

$$L^R = \{w^R \mid w \in L\},$$

the *reversal language of L* (where w^R denotes the reversal of the string w). Prove that if L is context-free, then L^R is also context-free.

“B problems” must be turned in.

Problem B1 (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 $M: \Sigma^* \rightarrow 2^{\Delta^*}$. 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 output 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 $M: \Sigma^* \rightarrow 2^{\Delta^*}$ 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 $f: T^* \rightarrow \Sigma^*$ and $g: T^* \rightarrow \Delta^*$ 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).

Problem B2 (80 pts). Let $D = (Q, \Sigma, \delta, q_0, F)$ be a DFA with n states, say q_1, \dots, q_n , where q_1 is the start state (Note, we denote the start state q_1 , not q_0 !) We associate with D the $n \times n$ boolean matrix, Δ_D , defined such that

$$\Delta_D(q_i, q_j) = \begin{cases} 1 & \text{if } (\exists a \in \Sigma)(\delta(q_i, a) = q_j), \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $\Delta_D(q_i, q_j) = 1$ iff there is some edge from q_i to q_j (regardless of the label of that edge). Add and multiply matrices treating $\{0, 1\}$ as truth values, i.e.

$$\begin{aligned} 0 + 0 &= 0 \\ 0 + 1 &= 1 \\ 1 + 0 &= 1 \\ 1 + 1 &= 1 \\ 00 &= 0 \\ 01 &= 0 \\ 10 &= 0 \\ 11 &= 1. \end{aligned}$$

(In other words, $\{0, 1\}$ is the two-element boolean ring).

(a) Prove that Δ_D^k gives the k -step reachability relation on D , i.e., $\Delta_D^k(q_i, q_j) = 1$ iff $\delta^*(q_i, w) = q_j$, for some string $w \in \Sigma^*$ with $|w| = k$ (We set $\Delta_D^0 = I$, the identity matrix).

Prove that there are only finitely many matrices Δ_D^k , where $k \geq 0$. For any $k \geq 0$, let

$$\Delta_D^{*[k]} = I + \Delta_D + \Delta_D^2 + \cdots + \Delta_D^k.$$

Prove that there is a smallest $k \leq n - 1$ so that $\Delta_D^{*[k]} = \Delta_D^{*[k+i]}$ for all $i \geq 1$. Let $\Delta_D^* = \Delta_D^{*[k]}$, for the above k . Prove that $\Delta_D^*(q_i, q_j) = 1$ iff $\delta^*(q_i, w) = q_j$, for some string $w \in \Sigma^*$.

(b) For simplicity of notation, drop the subscript D in Δ_D . Prove that there are only finitely many matrices of the form Δ^{2^n} . Use this fact to prove that for any regular language, L , the language

$$L_{2^n} = \{u \in \Sigma^* \mid (\exists v \in \Sigma^*)(|v| = 2^{|u|} \text{ and } uv \in L)\}$$

is also regular. Prove that the language

$$L_{2^{2^n}} = \{u \in \Sigma^* \mid (\exists v \in \Sigma^*)(|v| = 2^{2^{|u|}} \text{ and } uv \in L)\}$$

is also regular.

In answering this question, you do not have to be very specific about the details of your construction. Sketch clearly your construction without worrying too much about its actual implementation.

(c) Observe that $\Delta^{(n+1)^2} = \Delta^{n^2} \Delta^{2n} \Delta$. Prove that for any regular language, L , the language

$$L_{n^2} = \{u \in \Sigma^* \mid (\exists v \in \Sigma^*)(|v| = |u|^2 \text{ and } uv \in L)\}$$

is also regular.

Hint. Use the finite-state control to encode transitions between pairs of boolean matrices of the form

$$(C, D) \longrightarrow (CD\Delta, D\Delta^2).$$

Prove that

$$(I, I) \xrightarrow{n} (\Delta^{n^2}, \Delta^{2n}).$$

As in (b) you do not have to be very specific about the details of your construction.

(d) Let $P(n)$ be any polynomial with integer coefficients, say $P(n) = a_0n^d + a_1n^{d-1} + \dots + a_d$. Denote the i th derivative of $P(n)$ by $P^{(i)}(n)$ (where $i \geq 0$, with $P^{(0)}(n) = P(n)$). Prove that

$$P(n+1) = \sum_{i=0}^d \frac{P^{(i)}(n)}{i!}.$$

Hint. Prove it for $P(n) = n^d$ and use linearity.

Prove that

$$\frac{P^{(j)}(n+1)}{j!} = \sum_{k=j}^d \binom{k}{j} \frac{P^{(k)}(n)}{k!} \quad \text{and} \quad \frac{P^{(j)}(0)}{j!} = a_{d-j}.$$

Use all this to prove that for every regular language, L , and any polynomial, $P(n)$, the language

$$L_P = \{u \in \Sigma^* \mid (\exists v \in \Sigma^*)(|v| = P(|u|) \quad \text{and} \quad uv \in L)\}$$

is a regular language.

Hint. Encode in the finite-state control transitions between $(d+1)$ -tuples of boolean matrices (C_1, \dots, C_n) , where

$$(C_1, \dots, C_n) \longrightarrow (D_1, \dots, D_n) \quad \text{where} \quad D_j = \prod_{k=j}^d C_k^{\binom{k}{j}},$$

with $0 \leq j \leq d$. Then,

$$(\Delta^{a_{d-i}})_{i=0}^d \xrightarrow{n} (\Delta^{\frac{P^{(i)}(n)}{i!}})_{i=0}^d$$

and the first component on the righthand side is $\Delta^{P(n)}$.

As in (b) and (c) you do not have to be very specific about the details of your construction.

Problem B3 (150 pts). Review the definition of an ultimately periodic subset of \mathbb{N} from Homework I, Problem B5. A function, $f: \mathbb{N} \rightarrow \mathbb{N}$ *preserves ultimate periodicity* iff $f^{-1}(U)$ is ultimately periodic whenever $U \subseteq \mathbb{N}$ is ultimately periodic.

Given any language, $L \subseteq \Sigma^*$, let

$$\begin{aligned} L_f &= \{u \in \Sigma^* \mid (\exists v \in \Sigma^*)(|v| = f(|u|) \quad \text{and} \quad uv \in L)\} \\ L'_f &= \{u \in \Sigma^* \mid (\exists v \in \Sigma^*)(|v| = f(|u|) \quad \text{and} \quad v \in L)\}. \end{aligned}$$

(a) Assume that L is a regular language. Prove that for any function, $f: \mathbb{N} \rightarrow \mathbb{N}$, if f preserves ultimate periodicity then L'_f is regular.

(b) Prove that if L'_f is regular whenever L is regular, then L_f is regular whenever L is regular. Conclude that for any function, $f: \mathbb{N} \rightarrow \mathbb{N}$, if f preserves ultimate periodicity then L_f is regular.

(c) Prove that if L_f is regular whenever L is regular, then f preserves ultimate periodicity.

Therefore, L_f is regular whenever L is regular iff f preserves ultimate periodicity.

(d) Checking that a function preserves ultimate periodicity is usually difficult. Hence, it is desirable to find more convenient criteria for the preservation of ultimate periodicity. Such a criterion is given below.

For any two integers $m \geq 0$ and $p \geq 1$, let $[m]_p$ denote the set of integers congruent to m modulo p , i.e., $[m]_p = \{n \in \mathbb{N} \mid n = pi + m, i \in \mathbb{Z}\}$. (Note: \mathbb{Z} consists of the positive and negative integers, \mathbb{N} of the *non-negative* integers.) Check quickly that Homework I, Problem B5, asserts that every ultimately periodic subset, U , of \mathbb{N} can be written as

$$U = F \oplus \bigcup_{i=1}^k [m_i]_p,$$

for some $p \geq 1$ and some pairwise distinct m_i 's, with F a finite set disjoint from each of the $[m_i]_p$. Here $X \oplus Y = (X - Y) \cup (Y - X)$, the symmetric difference of X and Y , i.e., \oplus corresponds to *exclusive or*. Also, the set of ultimately periodic subsets of \mathbb{N} is the smallest family containing all the finite sets and the sets $[m]_p$ and closed under union, intersection and complementation. Check that if U and V are two (infinite) ultimately periodic sets with periods p_1 and p_2 , then $U \cup V$ is ultimately periodic with period $\text{lcm}(p_1, p_2)$.

We say that a function, $f: \mathbb{N} \rightarrow \mathbb{N}$, is *ultimately periodic modulo m* (with $m \geq 1$) iff there is some $p \geq 1$ and some $N \geq 0$ so that

$$f(n) \equiv f(n + p) \pmod{m}, \quad \text{for all } n \geq N.$$

Prove that a function, f , is ultimately periodic modulo m for all $m \geq 1$ iff $f^{-1}([i]_m)$ is ultimately periodic for all $i \geq 0$ and all $m \geq 1$.

Assume that the function, f , is ultimately periodic modulo m for all $m \geq 1$ and that $f^{-1}(\{n\})$ is ultimately periodic for every $n \in \mathbb{N}$. Then, prove that f preserves ultimate periodicity. Prove that f preserves ultimate periodicity iff

- (i) The function, f , is ultimately periodic modulo m for all $m \geq 1$, and
- (ii) The set $f^{-1}(\{n\})$ is ultimately periodic for every $n \in \mathbb{N}$.

Remark: Using the above, it is easy to show that the class of functions that preserve ultimate periodicity contains the functions, n^k , k^n (for any fixed $k > 1$), and is closed under composition, addition, multiplication and exponentiation.

(e) Reprove B2 using B3(b,d) (i.e., show that the languages L_{2^n} , $L_{2^{2^n}}$, L_{n^2} , L_P , are regular if L is regular).

(f) Prove that the function $f(n) = \log n$ does not preserve ultimate periodicity. Deduce that there are regular languages, L , for which

$$L_{\log n} = \{u \in \Sigma^* \mid (\exists v \in \Sigma^*)(|v| = \log(|u|) \text{ and } uv \in L)\}$$

is **not** regular.

(g) (**Extra Credit (40 pts).**) Prove again that if L is regular and f preserves ultimate periodicity, then L_f is regular, using Myhill-Nerode as suggested below.

If $L = L(D)$, for a trim DFA, D , the Myhill-Nerode equivalence relation \simeq_D induces n equivalence classes, L_1, \dots, L_n , which are all regular. Prove that for each L_i , the language, $D_u(L) = \{v \in \Sigma^* \mid uv \in L\}$, where $u \in L_i$, does not depend on the choice of u . Thus, for each L_i , let $R_i = D_{u_i}(L)$, where u_i is any string in L_i , for $i = 1, \dots, n$. Since $\Sigma^* = L_1 \cup \dots \cup L_n$, we can write

$$L_f = (L_f \cap L_1) \cup \dots \cup (L_f \cap L_n).$$

To prove that L_f is regular, it is enough to prove that each $L_f \cap L_i$ is regular. Show that

$$L_f \cap L_i = \{u \in L_i \mid (\exists v \in R_i)(|v| = f(|u|))\},$$

and finish up the proof (use Homework I, Problem B5).

Problem B4 (40 pts). The purpose of this problem is to get a fast algorithm for testing state equivalence in a DFA. Let $D = (Q, \Sigma, \delta, q_0, F)$ be a deterministic finite automaton. Recall that *state equivalence* is the equivalence relation \equiv on Q , defined such that,

$$p \equiv q \text{ iff } \forall z \in \Sigma^* (\delta^*(p, z) \in F \text{ iff } \delta^*(q, z) \in F).$$

and that *i-equivalence* is the equivalence relation \equiv_i on Q , defined such that,

$$p \equiv_i q \text{ iff } \forall z \in \Sigma^*, |z| \leq i (\delta^*(p, z) \in F \text{ iff } \delta^*(q, z) \in F).$$

A relation $S \subseteq Q \times Q$ is a *forward closure* iff it is an equivalence relation and whenever $(p, q) \in S$, then $(\delta(p, a), \delta(q, a)) \in S$, for all $a \in \Sigma$.

We say that a forward closure S is *good* iff whenever $(p, q) \in S$, then *good*(p, q), where *good*(p, q) holds iff either both $p, q \in F$, or both $p, q \notin F$.

Given any relation $R \subseteq Q \times Q$, recall that the smallest equivalence relation R_{\approx} containing R is the relation $(R \cup R^{-1})^*$ (where $R^{-1} = \{(q, p) \mid (p, q) \in R\}$, and $(R \cup R^{-1})^*$ is the reflexive

and transitive closure of $(R \cup R^{-1})$. We define the sequence of relations $R_i \subseteq Q \times Q$ as follows:

$$R_0 = R_{\approx}$$
$$R_{i+1} = (R_i \cup \{(\delta(p, a), \delta(q, a)) \mid (p, q) \in R_i, a \in \Sigma\})_{\approx}.$$

(i) Prove that $R_{i_0+1} = R_{i_0}$ for some least i_0 . Prove that R_{i_0} is the smallest forward closure containing R .

We denote the smallest forward closure R_{i_0} containing R as R^\dagger , and call it the *forward closure of R* .

(ii) Prove that $p \equiv q$ iff the forward closure R^\dagger of the relation $R = \{(p, q)\}$ is good.

TOTAL: 320 points.