$\boxed{\textbf{Fall 2021} \quad \textbf{CIS 511}}$

Introduction to the Theory of Computation
Jean Gallier

Homework 1

September 1, 2021; Due September 27, 2021

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

Problem A1. Let Σ be an alphabet, for any languages $L_1, L_2, L_3 \subseteq \Sigma^*$, prove that if $L_1 \subseteq L_2$, then $L_1L_3 \subseteq L_2L_3$.

Problem A2. Let Σ be an alphabet. Given any two families of languages $(A_i)_{i\in I}$ and $(B_j)_{j\in J}$, where I and J are any arbitrary index sets and $A_i, B_j \subseteq \Sigma^*$, prove that

$$\left(\bigcup_{i\in I} A_i\right)\left(\bigcup_{j\in J} B_j\right) = \bigcup_{(i,j)\in I\times J} A_iB_j.$$

"B problems" must be turned in.

Problem B1 (20 pts). Given an alphabet Σ, for any language $L \subseteq \Sigma^*$, prove that $L^*L^* = L^*$ and $L^{**} = L^*$.

Hint. To prove that $L^{**} = L^*$, prove that $(L^*)^n = L^*$ for all $n \geq 1$.

Problem B2 (70 pts). Let $\Sigma = \{a_1, \ldots, a_k\}$ be any alphabet. Given a string $w \in \Sigma^*$, its reversal w^R is defined inductively as follows: $\epsilon^R = \epsilon$, and $(ua)^R = au^R$, where $a \in \Sigma$ and $u \in \Sigma^*$.

A *palindrome* is a string w such that $w = w^R$. Here are some examples of palindromes:

eye
racecar
never odd or even
god saw I was dog
campus motto bottoms up mac
do geese see god

If $k = 1$, every string is a palindrome. Therefore we assume that $k \geq 2$.

We would like to give a formula giving the number p_n of all palindromes w of length $|w| = n \geq 0$ over the alphabet $\Sigma = \{a_1, \ldots, a_k\}$ with k letters.

(1) Prove that a palindrome $w \in \Sigma^*$ is either the empty string $w = \epsilon$, or $w = a$ with $a \in \Sigma$, or $w = aua$ where u is a palindrome of length $n - 2$ where $n = |w| \geq 2$ and $a \in \Sigma$ is some letter.

(2) Prove that $p_0 = 1$, $p_1 = k$, and

$$p_{n+2} = kp_n, \quad \text{for all } n \geq 0.$$

Give a formula for p_n. Distinguish between the cases where $n = 2m$ (n is even) and $n = 2m + 1$ (n is odd). You must prove the correctness of your formulae (use induction).

Do **not** give formulae in terms of $n/2$ when n is even or $(n - 1)/2$ when n odd. Please give formulae for p_{2m} and p_{2m+1} in terms of m.

(3) Prove that the number P_n of all palindromes w of length $\leq n$ (which means that $0 \leq |w| \leq n$) over the alphabet $\Sigma = \{a_1, \ldots, a_k\}$ with k letters is given by

$$\begin{aligned} P_{2m} &= \frac{2k^{m+1} - k - 1}{k - 1} & & n = 2m \\ P_{2m+1} &= \frac{k^{m+2} + k^{m+1} - k - 1}{k - 1} & & n = 2m + 1, \end{aligned}$$

for any natural number $m \in \mathbb{N}$. Prove that the number Q_n of all non-palindromes w of length $\leq n$ over the alphabet $\Sigma = \{a_1, \ldots, a_k\}$ is given by

$$\begin{aligned} Q_{2m} &= \frac{k^{2m+1} - 2k^{m+1} + k}{k - 1} & & n = 2m \\ Q_{2m+1} &= \frac{k^{2m+2} - k^{m+2} - k^{m+1} + k}{k - 1} & & n = 2m + 1, \end{aligned}$$

for any natural number $m \in \mathbb{N}$.
Hint. Figure out the total number of strings of length $\leq n$ over an alphabet of size $k \geq 2$.

(4) If $k = 2$, prove that if $m \geq 2$, then $P_{2m}/Q_{2m} < 1$ and $P_{2m+1}/Q_{2m+1} < 1$, so there are more non-palindromes than palindromes. What is 536 870 909 (in relation to palindromes)? Show that

$$\frac{536\,870\,909}{2^{55} - 1} \approx 2^{-26} \approx 1.4901 \times 10^{-8}.$$

What the interpretation of the above ratio as a probability?

Problem B3 (30 pts). Let Σ be any alphabet. For any string $w \in \Sigma^*$ recall that w^n is defined inductively as follows:

$$\begin{aligned} w^0 &= \epsilon \\ w^{n+1} &= w^n w, \quad n \in \mathbb{N}. \end{aligned}$$

Prove the following property: for any two strings $u, v \in \Sigma^*$, $uv = vu$ iff there is some $w \in \Sigma^*$ such that $u = w^m$ and $v = w^n$, for some $m, n \geq 0$.

Hint. In the "hard" direction, consider the subcases

(1) $|u| = |v|$,

(2) $|u| < |v|$ and

(3) $|u| > |v|$

and use an induction on $|u| + |v|$.

Problem B4 (30 pts). For any language $L \subseteq \{a\}^*$, prove that if $L = L^*$, then there is a finite language $S \subseteq L$ such that $L = S^*$.

Hint. If $L \neq \{\epsilon\}$, then L contains some nonempty string, and there is a shortest nonempty string $a^m \in L$. Consider the finite set S of strings in L of the form a^{mq+r}, where $0 \leq r \leq m-1$, and where $q \geq 1$ is minimal.

Problem B5 (60 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 Definition 3.5 of the notes). 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 B6 (70 pts). In this problem, all DFA's under consideration use the same alphabet Σ.

(a) Given any two DFA's D_1 and D_2, prove that there is a DFA D and two DFA F-maps $\pi_1\colon D \to D_1$ and $\pi_2\colon D \to D_2$ such that the following *universal mapping property of products* holds: For any DFA M and any two DFA F-maps $f\colon M \to D_1$ and $g\colon M \to D_2$, there is a *unique* DFA F-map $h\colon M \to D$ such that

$$f = \pi_1 \circ h \quad \text{and} \quad g = \pi_2 \circ h,$$

as shown in the diagram below:

$$\begin{tikzcd} & D_1 \\ M \arrow[ur, "f"] \arrow[r, "h"] \arrow[dr, "g"'] & D \arrow[u, "\pi_1"'] \arrow[d, "\pi_2"] \\ & D_2 \end{tikzcd}$$

Moreover, prove that π_1 and π_2 are surjective. Prove that D is unique up to a DFA F-map isomorphism. This means that if D' is another DFA and if there are two DFA F-maps $\pi_1' \colon D' \to D_1$ and $\pi_2' \colon D' \to D_2$ such that the universal mapping property of products holds, then there are two unique DFA F-maps $\varphi \colon D \to D'$ and $\varphi' \colon D' \to D$ so that $\varphi' \circ \varphi = \mathrm{id}_D$, $\pi_1 = \pi_1' \circ \varphi$, $\pi_2 = \pi_2' \circ \varphi$, $\varphi \circ \varphi' = \mathrm{id}_{D'}$, $\pi_1' = \pi_1 \circ \varphi'$ and $\pi_2' = \pi_2 \circ \varphi'$. What is the language accepted by D?

Remark: We call D the *product of* D_1 *and* D_2 and we denote it by $D_1 \prod D_2$.

(b) Given any three DFA's D_1, D_2, and D_3 and any two DFA F-maps $f \colon D_1 \to D_3$ and $g \colon D_2 \to D_3$, prove that there is a DFA D and two DFA F-maps $\pi_1 \colon D \to D_1$ and $\pi_2 \colon D \to D_2$ such that

$$f \circ \pi_1 = g \circ \pi_2,$$

as in the diagram below

$$\begin{tikzcd} D \arrow[r, "\pi_1"] \arrow[d, "\pi_2"'] & D_1 \arrow[d, "f"] \\ D_2 \arrow[r, "g"'] & D_3 \end{tikzcd}$$

and the following *universal mapping property of fibred products* holds: for any DFA M and any two DFA F-maps $f' \colon M \to D_1$ and $g' \colon M \to D_2$ such that

$$f \circ f' = g \circ g',$$

as in the diagram below

$$\begin{tikzcd} M \arrow[r, "f'"] \arrow[d, "g'"'] & D_1 \arrow[d, "f"] \\ D_2 \arrow[r, "g"'] & D_3 \end{tikzcd}$$

there is a *unique* DFA F-map $h \colon M \to D$ such that

$$f' = \pi_1 \circ h \quad \text{and} \quad g' = \pi_2 \circ h,$$

as in the diagram below

$$\begin{tikzcd}
M \arrow[dr, dotted, "h"] \arrow[drr, "f'"] \arrow[ddr, "g'"'] & & \\
& D \arrow[r, "\pi_1"'] \arrow[d, "\pi_2"] & D_1 \arrow[d, "f"] \\
& D_2 \arrow[r, "g"'] & D_3
\end{tikzcd}$$

Prove that D is unique up to a DFA F-map isomorphism. This means that if D' is another DFA and if there are two DFA F-maps $\pi_1'\colon D' \to D_1$ and $\pi_2'\colon D' \to D_2$ such that

$$f \circ \pi_1' = g \circ \pi_2'$$

and the universal mapping property of fibred products holds, then there are two unique DFA F-maps $\varphi\colon D \to D'$ and $\varphi'\colon D' \to D$ so that $\varphi' \circ \varphi = \mathrm{id}_D$ $\pi_1 = \pi_1' \circ \varphi$, $\pi_2 = \pi_2' \circ \varphi$, $\varphi \circ \varphi' = \mathrm{id}_{D'}$, $\pi_1' = \pi_1 \circ \varphi'$ and $\pi_2' = \pi_2 \circ \varphi'$.

Remark: We denote D by $D_1 \prod_{D_3} D_2$ and call it a *fibred product of* D_1 *and* D_2 *over* D_3, or a *pullback of* D_1 *and* D_2 *over* D_3.

If T is any one-state DFA accepting Σ^* (this single state is accepting), observe that there is a unique DFA F-map from every DFA D to T. Use this to show that if $D_1 \prod D_2$ is the product DFA arising in (a), then

$$D_1 \prod D_2 = D_1 \prod_T D_2.$$

Extra Credit (40 points). Redo questions (a) and (b) for B-maps instead of F-maps.

Remark: If we dualize (b), i.e., turn the arrows around, we get the notion of *fibred coproduct* or *pushout*. It can be shown that fibred coproducts exist, both for F-maps and B-maps, but this is tricky.

TOTAL: 280 points + 40 extra credit.