CS 374 Problem Set 1

Problem 1

1.

Define function h mapping regular expressions over Σ to regular expressions over Δ as follows:

$$h(r) = \begin{cases} \emptyset & \text{if } r = \emptyset \\ \epsilon & \text{if } r = \epsilon \\ h(a) & \text{if } r = a \in \Sigma \\ (h(r_1) + h(r_2)) & \text{if } r = (r_1 + r_2) \\ (h(r_1)h(r_2)) & \text{if } r = (r_1 r_2) \\ (h(r_1)^*) & \text{if } r = (r_1^*) \end{cases}$$

2.

Given property: For every pair of strings u, v, $h(uv) = h(u)h(v)$.

Let us first prove two lemmas that we will need for this proof:

Lemma 1: $f(A) \cup f(B) = f(A \cup B)$
Proof: Note the preimages $f^{-1}(A) = \{a \mid a \in A\}$ and $f^{-1}(B) = \{b \mid b \in B\}$. We also see that the preimage $f^{-1}(A \cup B) = \{x \mid x \in (A \cup B)\}$. Trivially, we see that $\{a \mid a \in A\} \cup \{b \mid b \in B\} = \{x \mid (x \in A) \cup (x \in B)\} = \{x \mid x \in (A \cup B)\}$. Since the function is applied to identical preimages, the images of $f(A) \cup f(B)$ and $f(A \cup B)$ must be equal.

$\square$

Lemma 2: $h(w^k) = h(w)^k \; \forall \, k \in \mathbb{N}$.
Proof: Prove by induction on k.
Base: $k = 0, \;\; h(w^0) = h(\epsilon) = \epsilon = h(w)^0$. Base case holds.
Inductive: Assume that $\forall \, k < n, \; h(w^k) = h(w)^k$.
Take $k = n$. $w^n = w \cdot w^{n-1} \Rightarrow h(w) = h(w \cdot w^{n-1}) = h(w)h(w^{n-1})$ by given property. By the inductive assumption, we see that $h(w)h(w^{n-1}) = h(w)(h(w))^{n-1} = (h(w))^n$.

$\square$

Proof will be completed by induction on the number of operations in regular expression r. Let us define function:

$$\#_{ops} : \; \{r \mid \text{r is a regular expression}\} \to \mathbb{N}$$

such that $\#_{ops}(r_1) =$ number of operations in regular expression r_1.

Base: $\#_{ops}(r) = 1$. This presents 3 cases: $r \in \{\emptyset,\ \epsilon,\ h(a)\}$.

- **Case** $\emptyset$: $r = \emptyset \Rightarrow L(r) = \emptyset \Rightarrow h(L(r)) = \emptyset = L(h(r))$
- **Case** ϵ: $r = \epsilon \Rightarrow L(r) = \{\epsilon\} \Rightarrow h(L(r)) = \{\epsilon\} = L(h(r))$
- **Case** $a \in \Sigma$: $r = a \Rightarrow L(r) = \{a\} \Rightarrow h(L(r)) = \{h(a)\} = L(h(r))$

Inductive Assumption: Assume that for all regular expressions with fewer than n operations, that $L(h(r)) = h(L(r))$.

A regular expression r with n operations can be written in one of 3 ways:

1. $r_1 + r_2$, where $\#_{ops}(r_1) = n - 1$ and $\#_{ops}(r_2) = 1$.
2. $r_1 r_2$, where $\#_{ops}(r_1) = n - 1$ and $\#_{ops}(r_2) = 1$.
3. (r_1^*), where $\#_{ops}(r_1) = n - 1$.

<u>**Case** $r = r_1 + r_2$:</u>

$$L(r_1 + r_2) = \{w \mid w \in (L(r_1) \cup L(r_2))\}$$

Since both $\#_{ops}(r_1) = n - 1 < n$ and $\#_{ops}(r_2) = 1 < n$, bot r_1 and r_2 fall under the inductive assumption and we can say that $h(L(r_1)) = L(h(r_2))$ and $h(L(r_2)) = L(h(r_2))$. Note that $L(r_1 + r_2) = L(r_1) \cup L(r_2)$ by definition.We can apply h to both sides and we get $h(L(r_1 + r_2)) = h(L(r_1)) \cup h(L(r_2)) = L(h(r_1)) \cup L(h(r_2))$. By Lemma #1, we see that:

$$L(h(r-1)) \cup L(h(r_2)) = L(h(r_1) + h(r_2)) = L(h(r_1 + r_2)) = L(h(r)) = h(L(r_1 + r_2)) = h(L(r))$$

<u>**Case** $r = r_1 r_2$:</u>

$$L(r_1 r_2) = \{w_1 w_2 \mid w_1 \in L(r_1),\ w_2 \in L(r_2)\} \Rightarrow h(L(r_1 r_2)) = \{h(w_1 w_2) \mid w_1 \in L(r_1),\ w_2 \in L(r_2)\}$$

By the given property, we can see that:

$$\{h(w_1 w_2) \mid w_1 \in L(r_1),\ w_2 \in L(r_2)\} = \{h(w_1)h(w_2) \mid w_1 \in L(r_1),\ w_2 \in L(r_2)\}$$

$$\{h(w_1)h(w_2) \mid w_1 \in L(r_1),\ w_2 \in L(r_2)\} = h(L(r_1)) \cdot h(L(r_2))$$

Since $\#_{ops}(r_1) = n - 1 < n$ and $\#_{ops}(r_2) = 1 < n$, the inductive assumption implies that:

$$h(L(r_1)) \cdot h(L(r_2)) = L(h(r_1)) \cdot L(h(r_2)) = L(h(r_1)h(r_2)) = L(h(r))$$

<u>**Case** (r_1^*):</u>

By definition, we know that $L(h(r_1^*)) = L((h(r_1))^*) = \{h(w)^k \mid w \in L(r_1),\ k \in \mathbb{N}\}$.

$$\{h(w)^k \mid w \in L(r_1),\ k \in \mathbb{N}\} = \{h(w^k) \mid w \in L(r_1),\ k \in \mathbb{N}\} \text{ by Lemma \#2.}$$

$$h(L(r_1)) = \{h(w) \mid w \in L(r_1)\} \Rightarrow h(L(r_1^*)) = \{h(w^k) \mid w \in L(r_1),\ k \in \mathbb{N}\}$$

the two sets are equal, therefore $L(h(r)) = L(h(r_1^*)) = h(L(r_1^*)) = h(L(r))$.

$\square$