HW1

Due: n/a

Reminder: Submit your assignment on Gradescope by the due date. Submissions must be typeset. Each page should include work for only one problem (i.e., make a new page/new pages for each problem). See the course syllabus for the late policy.

While collaboration is encouraged, please remember not to take away notes from any labs or collaboration sessions. Your work should be your own. Use of other third-party resources is strictly forbidden.

Please monitor Ed discussion, as we will post clarifications of questions there.

Problem 1

Prove that for any integer $k > 0$, there is a language L such that

(a) L can be recognized by a DFA with k states

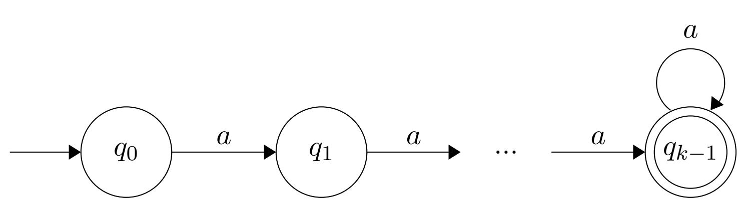
(b) No DFA with less than k states recognizes L

Solution:

Consider the language $L = \{w \mid w \text{ has length at least } k-1\}$ over alphabet $\{a\}$.

To show part (a), we first find a DFA that recognizes L that has exactly k states.

Consider the following DFA, D:



D has exactly K states, $q_0, \ldots, q_{k-1}$. If $w \notin L$, there must be less than $k-1$ symbols. Then, D must terminate in a state other than q_{k-1} and reject

w since q_{k-1} is the only accepting state. If $w \in L$, then w has length at least $k-1$. Then, D will reach q_{k-1} at some point and remain there since all transitions from q_{k-1} stay in q_{k-1}. This means that D will accept w. Therefore, $L(D) = L$.

We will show part (b) by contradiction. Assume that such a DFA, D, exists such that $L(D) = L$ and D has $k-1$ or less states. Consider the string w, which is a repeated $k-1$ times. w has length $k-1$ so $w \in L$ and is therefore accepted by D. D has less than $k-1$ states so by the pigeonhole principle some state must be visited at least twice. This means that D contains a loop and therefore some substring of w can be removed and the resulting string will still be accepted by D, meaning that $w \in L$. However, such a string would have length less than $k-1$ and so is not in L, a contradiction.

Problem 2

Design a DFA that recognizes the following language over the alphabet $\{0, 1\}$. Provide a proof of correctness.

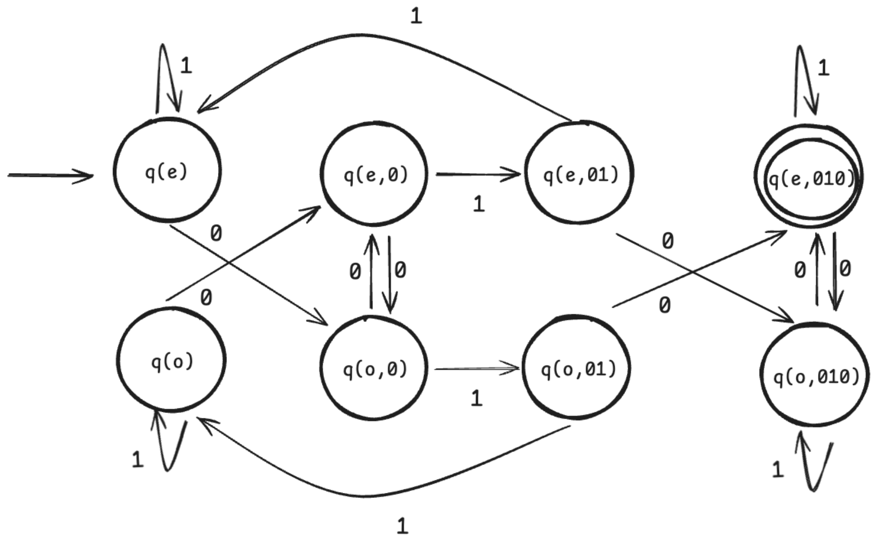
$\{w|\ w$ contains an even number of 0s and the substring 010$\}$

Solution:

Sample proof: design a DFA that recognizes

$$L = \{w \mid w \text{ contains an even number of 0s and the substring 010}\}$$

over the alphabet $\{0, 1\}$.



Proof of Correctness. We will show that $L(M) = L$.

Assume $w \in L$. Then w contains an even number of 0s and the substring 010. As M processes w, it tracks both whether the 0 count is even or odd (e or o) and whether the substring 010 has occurred yet. Since w contains 010, at some point M will reach state $q_{e,010}$ or $q_{o,010}$. Since w has an even number of 0s total, M will end in $q_{e,010}$ (an accepting state) after processing all of w.

On the other hand, assume $w \in L(M)$. Then M accepts w, meaning w causes M to end in an accepting state. The only accepting state is $q_{e,010}$, which is only reachable once the substring 010 has been seen and if the count of 0s so far is even. Thus, $w \in L$.

Problem 3

Let Σ and Δ be alphabets. Let h be a function that maps from Σ to Δ. Let $H : \Sigma^*$ to Δ^* be defined as follows on the string $w = w_1...w_n, w_i \in \Sigma$:

$$H(\epsilon) = \epsilon$$

$$H(w_1...w_n) = h(w_1)...h(w_n)$$

Let $L \subseteq \Sigma^*$. Any function H defined in this way is called a **homomorphism**. Let the image of L under H be:

$$H[L] = \{w \mid w = H(w_L) \text{ for some } w_L \in L\}$$

Let the inverse image of a language $L' \subseteq \Delta^*$ under H be:

$$H^{-1}[L'] = \{w \mid H(w) \in L'\}$$

1. Show that, if L is regular, $H[L]$ is regular.
2. Show that, if $L' \subseteq \Delta^*$ is regular, $H^{-1}[L']$ is regular. Hint: start from a DFA M accepting L', and construct M' such that it leverages M to recognize $H^{-1}[L']$

Solution:

1. *Proof that if L is regular, then $H[L]$ is regular.*

   If L is regular, then there exists a DFA D such that $L(D) = L$.

   We'll construct an NFA D' as follows: replicate all the states in D, but for every transition triggered by $w_i \in \Sigma^*$, replace this transition to be triggered by $h(w_i) \in \Delta^*$. This machine is non-deterministic because for two different symbols $w_1, w_2 \in \Sigma*$, we might have $h(w_1) = h(w_2) = w'_j$, and if there are different transitions specified by w_1 vs. w_2 from the same state in D, D' will simulate both of these transitions on input w_j.

   Proof of correctness (proof that $H[L] = L(D')$:

   Direction 1: Assume $w' = w'_1...w'_n \in H[L]$. Then there exists some $w \in L$ such that $H(w) = w'$. This means that, for each w'_i, there exists some w_i such that $h(w_i) = w'_i$ (by the definition of the homomorphism). We know that D terminates on accept for $w_1...w_n$, passing through states $q_1, ..., q_n$. But by the construction of D', it also has states $q_1, ..., q_n$, with transitions specified by $w'_1, ..., w'_n$. Therefore, D' accepts w', or $w' \in L(D')$.

   Direction 2: Assume $w' = w'_1...w'_n \in L(D')$, or that w' is accepted by D'. Then there was some path $q'_1, ..., q'_n$ in D' terminating on accept given input w'. By construction, we know there is also $q_1, ..., q_n$ in D, terminating in accept, where each transition is triggered by w_i such that $h(w_i) = w'_i$. Thus, $H(w) = w'$ for some $w \in L$, so $w' \in H[L]$.

2. *Proof that if $L' \subseteq \Delta^*$ is regular, then $H^{-1}[L']$ is regular.*

   If L' is regular, then there exists a DFA M such that $L(M) = L'$.

   We'll construct a DFA M' as follows: M' has the same state set as M, same start state, and same accept states. For the transition function, define $\delta'(q, \sigma) = \delta(q, h(\sigma))$ for all $q \in Q$ and $\sigma \in \Sigma$. This means M' simulates M by applying the homomorphism h to each input symbol before processing it.

   Proof of correctness (proof that $H^{-1}[L'] = L(M')$):

   Direction 1: Assume $w = w_1...w_n \in H^{-1}[L']$. Then by definition, $H(w) = h(w_1)...h(w_n) \in L' = L(M)$, so M accepts $H(w)$, passing through states $q_0, q_1, ..., q_n$ where q_n is an accept state. By construction of M', when M' processes w, it follows transitions $\delta'(q_{i-1}, w_i) = \delta(q_{i-1}, h(w_i)) = q_i$ for each i. This means M' follows exactly the same sequence of states $q_0, q_1, ..., q_n$ and terminates in the same accept state. Therefore, M' accepts w, or $w \in L(M')$.

   Direction 2: Assume $w = w_1...w_n \in L(M')$, or that w is accepted by M'. Then there was some path $q_0, q_1, ..., q_n$ in M' terminating on accept given input w, where $\delta'(q_{i-1}, w_i) = q_i$ for each transition. By construction of M', we know $\delta'(q_{i-1}, w_i) = \delta(q_{i-1}, h(w_i))$, so M also has the path $q_0, q_1, ..., q_n$ terminating in accept, triggered by input $h(w_1)...h(w_n) = H(w)$. Thus, M accepts $H(w)$, so $H(w) \in L(M) = L'$. Therefore $w \in H^{-1}[L']$ by definition.

Problem 4

For any string $w = w_1 w_2 \ldots w_n$, the reverse of w, written w^R, is the string w in reverse order: $w_n \ldots w_2 w_1$. For any language A, let $A^R = w^R | w \in A$. Show that if A is regular, so is A^R.

Solution:

Proof. Let A be regular, and let D be a DFA such that $L(D) = A$. We construct NFA N where:

- The start state, q_{start}, has an epsilon transition to each accept state in D
- N's accept state is D's start state, q_0.
- We invert the direction of each transition between states, from D to N.

Proof of correctness (proof that $A^R = L(N)$):

Direction 1: Assume $w = w_1...w_n \in A^R$. Then $w^R = w_n...w_1 \in A$. D accepts w with some path $q_0, q_1, ..., q_n$. By construction of N, there is an epsilon transition from q_{start} to q_n, and reversed transitions $q_n \xrightarrow{w_1} q_{n-1} \xrightarrow{w_2} q_{n-2} \cdots \xrightarrow{w_n} q_0$. Therefore, N has an accepting path from q_{start} to q_0 on input w, so N accepts w, meaning $w \in L(N)$.

Direction 2: Assume $w = w_1...w_n \in L(N)$, or that w is accepted by N. Then there exists an accepting path in N from q_{new} to q_0 on input w. This path starts with an epsilon transition to some q_f, followed by transitions $q_f, q_{f-1}, ...q_0$. By our construction of N, these correspond to reversed transitions in D: $q_0, ..., q_f$. Thus, D accepts the reverse of w: $w' = w_n...w_1$. This means $w = w'^R$, where $w' \in A$. So, $w \in A^R$.