Exam 1 solutions

1. This question consists of four parts, each of which asks you to prove a simple fact. Give a short proof (perhaps one or two sentences and/or formulas) for each fact.

(a) Let $A$ and $B$ be nonempty sets such that $A \subseteq B$. Prove that there exists an onto function of the form $f : B \rightarrow A$. (Consequently, any subset of a countable set is countable. You don’t have to argue this part...just prove that there exists an onto function of the form $f : B \rightarrow A$.)

**Solution.** Because $A$ is nonempty, we may fix an arbitrarily chosen element $a_0 \in A$. The function $f : B \rightarrow A$ defined as

$$f(b) = \begin{cases} b & \text{if } b \in A \\ a_0 & \text{if } b \notin A \end{cases}$$

for every $b \in B$ is an onto function of the required form.

(b) Prove that if $R_1$ and $R_2$ are regular expressions over some alphabet $\Sigma$, then there exists a regular expression $S$ such that $L(S) = L(R_1) \cap L(R_2)$.

**Solution.** Every language matched by a regular expression is regular, and therefore $L(R_1)$ and $L(R_2)$ are regular. The regular languages are closed under intersection (as discussed in lecture), and therefore $L(R_1) \cap L(R_2)$ is regular. Finally, every regular language is matched by some regular expression, and therefore there exists a regular expression $S$ such that $L(S) = L(R_1) \cap L(R_2)$.

(c) Let us say that a string $x$ is obtained from a string $w$ by deleting symbols if it is possible to remove zero or more symbols from $w$ so that just the string $x$ remains. For example, the following strings can all be obtained from 0110 by deleting symbols: $\varepsilon$, 0, 1, 00, 01, 10, 11, 010, 011, 110, and 0110.

Let $\Sigma = \{0, 1\}$, let $A \subseteq \Sigma^*$ be a regular language, and define

$$B = \left\{ x \in \Sigma^* : \text{there exists a string } w \in A \text{ such that } x \text{ is obtained from } w \text{ by deleting symbols} \right\}.$$

Prove that the language $B$ is regular.

**Solution.** Let $M = (Q, \Sigma, \delta, q_0, F)$ be a DFA such that $L(M) = A$. Define an NFA $N = (Q, \Sigma, \eta, q_0, F)$ as follows:

$$\eta(q, a) = \{ \delta(q, a) \} \quad \text{and} \quad \eta(q, \varepsilon) = \{ \delta(q, b) : b \in \Sigma \} \quad \text{(for all } q \in Q \text{ and } a \in \Sigma).$$

In words, $N$ is similar to $M$, but for every transition in $M$ we include in $N$ the same transition as well as an $\varepsilon$-transition between the same pair of states.

It is the case that $L(N) = B$, and therefore $B$ is regular.
We have already proved that the language $A = \{0^n1^n : n \in \mathbb{N}\}$ is nonregular. Prove that the language $A^*$ is also nonregular. (You can use the pumping lemma for this question if you wish, but there is an easier way.)

**Solution.** Assume toward contradiction that $A^*$ is regular. The language $L(0^*1^*)$ is regular, because it is the language matched by a regular expression, and therefore $A^* \cap L(0^*1^*)$ is also regular, because the regular languages are closed under intersection. However, it is the case that $A^* \cap L(0^*1^*) = A$, which contradicts the fact that $A$ is nonregular. Having obtained a contradiction, we conclude that $A^*$ is nonregular.

2. Let $\Sigma$ be an alphabet, let $A \subseteq \Sigma^*$ be a regular language, and let $B \subseteq \Sigma^*$ be an arbitrary language. Define a language $C \subseteq \Sigma^*$ as follows:

$$C = \{ w \in \Sigma^* : \text{there exists } x \in B \text{ such that } wx \in A \}.$$ 

In words, $C$ is the language of all strings that can be obtained by first choosing a string from $A$ and then removing from that string any suffix that is contained in $B$.

Prove that $C$ is regular.

Hint: do not be scared by the fact that $B$ is an arbitrary language. For any DFA $M = (Q, \Sigma, \delta, q_0, F)$, and any pair of states $p, q \in Q$, there either exists a string $x \in B$ such that $\delta^*(p, x) = q$, or there does not.

**Solution.** Let $M = (Q, \Sigma, \delta, q_0, F)$ be a DFA such that $L(M) = A$. Define a subset of states $G = \{ q \in Q : \delta^*(q, x) \in F \text{ for some } x \in B \}$,

and consider the DFA $K = (Q, \Sigma, \delta, q_0, G)$. It is the case that $L(K) = C$, and therefore $C$ is regular.

3. For any alphabet $\Sigma$ and any string $w \in \Sigma^*$, define $\text{rotate}(w) \in \Sigma^*$ to be the string that you would obtain by removing the rightmost symbol of $w$ and then adding that symbol to the left-hand side of $w$. In more precise terms, this operation is as follows:

(i) $\text{rotate}(\epsilon) = \epsilon$, and 
(ii) $\text{rotate}(wa) = aw$ for all $a \in \Sigma$ and $w \in \Sigma^*$.

For example, $\text{rotate}(000111) = 100111$.

Now suppose that $\Sigma = \{0, 1\}$ and $A \subseteq \Sigma^*$ is a regular language. Prove that the language

$$B = \{ \text{rotate}(w) : w \in A \}$$

is also regular.
**Solution.** Let $M = (Q, \Sigma, \delta, q_0, F)$ be a DFA such that $L(M) = A$. Define a DFA 

$$K = \left( (\Sigma \times Q) \cup \{r_0\}, \Sigma, \eta, r_0, G \right)$$

as follows. The transition function $\eta$ is defined as

$$\eta(r_0, 0) = (0, q_0)$$
$$\eta(r_0, 1) = (1, q_0)$$

and

$$\eta((a, q), b) = (a, \delta(q, b))$$

for all $q \in Q$ and $a, b \in \Sigma$. If it is the case that $\varepsilon \in A$, then we set

$$G = \{(a, q) : a \in \Sigma, \ q \in Q, \ \delta(q, a) \in F \} \cup \{r_0\},$$

and otherwise we set

$$G = \{(a, q) : a \in \Sigma, \ q \in Q, \ \delta(q, a) \in F \}.$$ 

Intuitively speaking, the DFA $K$ operates as follows. It remembers the first symbol that it reads, so that its state becomes $(a, q_0)$ after it reads the first symbol $a$. It then simulates $M$ on the remaining input symbols. The accept states are those states of the form $(a, q)$ for which the stored symbol $a$ would cause $M$ to transition from $q$ to an accept state. We have included $r_0$ in $G$ if $\varepsilon \in A$ as a special case, given that the question defines rotate($\varepsilon$) = $\varepsilon$.

It is the case that $L(K) = B$, and therefore $B$ is regular.

4. Let $\Sigma = \{0, 1\}$, and define a language

$$MIDDLE = \{u0v : u, v \in \Sigma^* \text{ and } |u| = |v|\}.$$ 

In words, MIDDLE is the language of all binary strings of odd length whose middle symbol is 0. Prove that MIDDLE is not regular.

In case you are inclined to use the pumping lemma when solving this problem, here it is:

**Pumping lemma.** For every alphabet $\Sigma$ and every regular language $A \subseteq \Sigma^*$, there exists a pumping length $n \geq 1$ for $A$ that satisfies the following property. For every string $w \in A$ with $|w| \geq n$, it is possible to write $w = xyz$ for strings $x, y, z \in \Sigma^*$ such that

1. $y \neq \varepsilon,$
2. $|xy| \leq n,$ and
3. $xy^iz \in A$ for every $i \in \mathbb{N}.$

**Solution.** Assume toward contradiction that MIDDLE is regular. By the pumping lemma, there exists a pumping length $n \geq 1$ for MIDDLE satisfying the property expressed in that lemma.
Define $w = 1^n01^n$. It is the case that $w \in \text{MIDDLE}$ and $|w| = 2n + 1 \geq n$. Therefore, by the pumping lemma, it is possible to write $w = xyz$ for strings $x, y, z \in \Sigma^*$ such that (i) $y \neq \epsilon$, (ii) $|xy| \leq n$, and (iii) $xy^iz \in \text{MIDDLE}$ for all $i \in \mathbb{N}$.

Because $w = xyz = 1^n01^n$ and $|xy| \leq n$, it must be the case that $y = 1^k$ for some natural number $k \in \mathbb{N}$, as the prefix $xy$ of $w$ is not long enough to reach the single 0 that string contains. Moreover, because $y \neq \epsilon$ we conclude that $k \geq 1$. Setting $i = 2$, we find from the third condition on $x, y, z$ that

$$xy^2z = xyyz = 1^{n+k}01^n \in \text{MIDDLE}.$$  

However, because $k \geq 1$, the string $1^{n+k}01^n$ is clearly not in MIDDLE; either its length is even (in case $k$ is odd) or it has a 1 rather than a 0 in its middle position (in case $k$ is even).

Having obtained a contradiction, we conclude that MIDDLE is not regular.