CSC236 Summer 2017
Assignment #3: Formal Languages and Finite Automata
Solutions
Due August 10th, by 6:00 pm

The aim of this assignment is to give you some practice in formal languages and finite automata.

For each question, please write up detailed answers carefully. Make sure that you use notation and terminology correctly, and that you explain and justify what you are doing. Marks will be deducted for incorrect or ambiguous use of notation and terminology, and for making incorrect, unjustified, ambiguous, or vague claims in your solutions. You should clearly cite any sources or people you consult, other than the course notes, lecture materials, and tutorial exercises.

Your assignment must be typed to produce a PDF document **A3.pdf** (hand-written submissions are not acceptable). You may work on the assignment in groups of 1, 2, or 3, and submit a single assignment for the entire group on MarkUs.

1. Give incomplete DFA's (not including the dead states) accepting the following languages over the alphabet $\{0, 1\}$:

   (a) The set of all strings with three consecutive 1's (these instances of 1's can appear at any part of the string).

   (b) The set of all strings with the 101 as a substring.

   (c) The set of strings ending in 11.

Answer

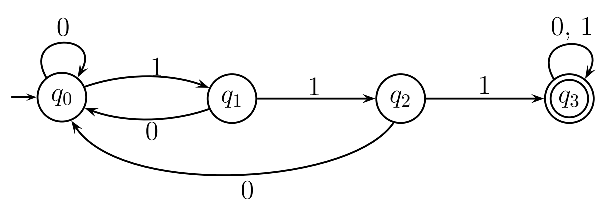

Figure 1: A DFA accepting the language asked in 1.(a).

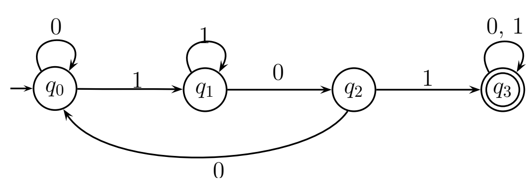

Figure 2: A DFA accepting the language asked in 1.(b).

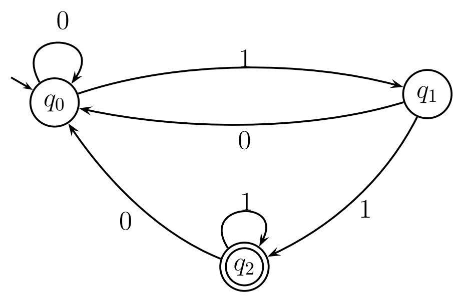

Figure 3: A DFA accepting the language asked in 1.(c).

2. Let A be a DFA and q a particular state of A, such that $\delta(q, a) = q$ for all input symbols a. Show by induction on the length of the input that for all input strings w, we have $(q, w) \vdash^* (q, \varepsilon)$.

Answer

We use mathematical induction.

Base Case: For $|w| = 0$, $(q, w) \vdash (q, \varepsilon)$.
Induction Hypothesis: $(q, w) \vdash (q, \varepsilon)$, where $0 \leq |w| \leq k$ for some k.
Inductive Step: For $|w| = k + 1$, let us denote $w = \sigma w_0$ where $|w_0| = k$. From inductive hypothesis, $(q, w_0) \vdash (q, \varepsilon)$. Since $\delta(q, a) = q$ for all input symbols a, $\delta(q, \sigma) = q$. Therefore, $(q, w) = (q, \sigma w_0) \vdash (q, w_0) \vdash^* (q, \varepsilon)$. So the statement holds for all input strings w.

3. Draw the state diagram of the following NFA and convert it to a DFA. Informally describe the language it accepts.

	0	1
$\rightarrow p$	$\{p, q\}$	$\{p\}$
q	$\{r, s\}$	$\{s\}$
r	$\{p, r\}$	$\{s\}$
$*s$	$\emptyset$	$\emptyset$

Answer

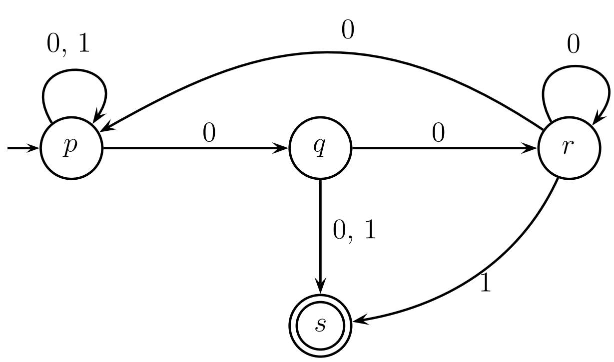

Figure 4: The NFA decribed in the above table.

The above finite automata accept all the stings over 0's and 1's which end with 00 or 01.

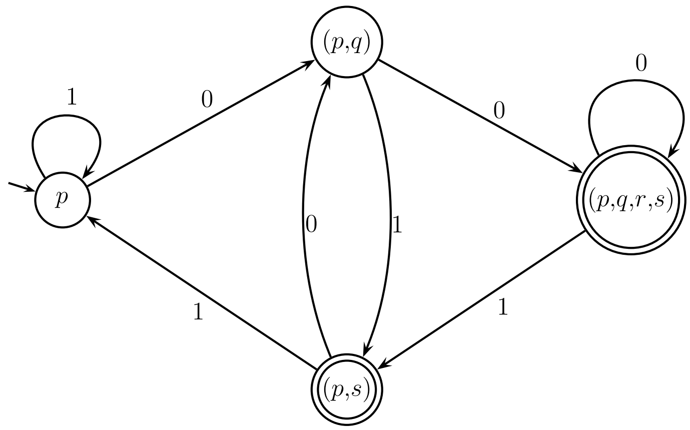

Figure 5: The resulting DFA that we get by applying the subset construction.

4. Write regular expressions for the following languages:

   (a) The set of strings of 0's and 1's whose seventh symbol from the right end is 1.

   (b) The set of 0's and 1's with at most one pair of consecutive 1's.

   (c) The set of all strings of 0's and 1's not containing 110 as a substring.

Answer

   (a) $(0+1)^*1(0+1)(0+1)(0+1)(0+1)(0+1)(0+1)$.

   (b) $(0+10)^*(11+1+\varepsilon)(0+01)^*$.

   (c) $(0+10)^*1^*$.

5. Prove or disprove each of the following statements about regular expressions.

   (a) $(R+S)^* = R^* + S^*$

   (b) $(RS+R)^*RS = (RR^*S)^*$

   (c) $(R+S)^* = (R^*S)^*$

   (d) $S(RS+S)^*R = RR^*S(RR^*S)^*$

Answer

All of the above statements are not true. Let for all of them R to be 0 and S to be 1, then we have the following:

   (a) $01 \in (0+1)^*$, but $01 \notin 0^* + 1^*$.

   (b) $\varepsilon \notin (01+0)^*01$, but $\varepsilon \in (00^*1)^*$.

   (c) $000 \in (0+1)^*$, but $000 \notin (0^*1)^*$.

   (d) $10 \in 1(01+1)^*0$, but $10 \notin 00^*1(00^*1)^*$.

6. The finite automaton if Figure 6 accepts no words of length zero or length one, it accepts only two words of length two (01 and 10). There is a recurrence equation for the number $N(k)$ of words of length k that this automaton accepts (i.e. $N(0) = 0$, $N(1) = 0$, $N(2) = 2$, $N(3) = 0$, $N(4) = 2$ etc.). Discover this recurrence and demonstrate your understanding by identifying the number $N(14)$ (give a short explanation not just a random number).

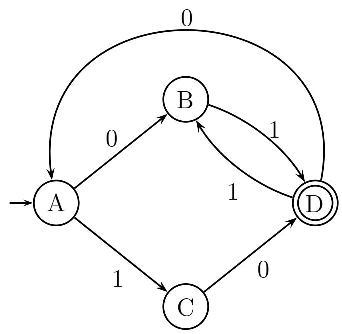

Figure 6: A DFA accepting a language over the alphabet $\{0, 1\}$.

Answer

We have to notice that every accepting word w, with length greater than two, has to finish with 11, 001, or 010. If the word $w = w_0 11$ is accepted, then the word w_0 is accepted. Similar, for $w = w_0 001$ or $w = w_0 010$. Then, we have that for all $k \geq 3$

$$N(k) = N(k-2) + 2 \cdot N(k-3)$$

We know already the first four values of the N function, $N(0) = 0$, $N(1) = 0$, $N(2) = 2$, $N(3) = 0$, and $N(4) = 2$. Then, we can continue as follows, $N(5) = 4$, $N(6) = 2$, $N(7) = 8$, $N(8) = 10$, $N(9) = 12$, $N(10) = 26$, $N(11) = 32$, and $N(12) = 50$. Hence, we have that

$$N(14) = 114$$

7. Prove that the following languages are not regular. You may use the pumping lemma and the closure properties of regular languages under union, intersection and complement.

   (a) $L = \{0^n 1^m 0^n \mid m, n \geq 0\}$.

   (b) $L = \{w \in \{0, 1\}^* \mid w \text{ is palindrome}\}$.
   (A palindrome is a string that reads the same forward and backward. e.g.: racecar)

Answer

(a) Suppose L is regular. Then there would be a pumping constant n and take the string $w = 0^n 1^n 0^n \in L$. By the pumping lemma, we can break $w = xyz$ such that $|y| > 0$ and $|xy| \leq n$. Since $|xy| \leq n$ we have that $x, y \in 0^*$. Now consider the string $xy^0 z$, then the string $0^{n-m} 1^n 0^n$ should also be in L, for $1 \leq m \leq n$, which is a contradiction. Therefore, L is not regular.

(b) Let $L_1 = \{0^n 1^m 0^n \mid m, n \geq 0\}$ which we just prove that it is not regular. Let also $L_2 = \{w \in \{0, 1\}^* \mid w \text{ is palindrome}\}$. We have that $L_1 = L_2 \cap L(0^* 1^* 0^*)$, which means that if L_2 is regular, so is $L_2 \cap L(0^* 1^* 0^*)$, but we have seen already that it is not. Then L cannot be regular.

8. Show that the regular languages are closed under the following operation.

$$\frac{1}{2}(L) = \{x \in \Sigma^* \mid \text{there exists } y \in \Sigma^* \text{ with } |y| = |x| \text{ such that } xy \in L\}.$$

Answer

Consider a DFA $A = (Q, \Sigma, \delta, q_0, F)$ such that $L(A) = L$. Then we can build a DFA $A' = (Q'\Sigma, \delta', q_0', F')$ as follows: $Q' = \{(q, S) \mid q \in Q, S \subseteq Q\}$, $q_0' = (q_0, F)$, $F' = \{(q, S) \in Q' \mid q \in S\}$, and δ' is defined by:

$$\delta'((q, S), a) = (\delta(q, a), T) \text{ where}$$

$$T = \{p \in Q \mid \exists \sigma \in \Sigma, \text{ such that } \delta(p, \sigma) \in S\}$$

We will argue now that $\frac{1}{2}(L) = L(A')$. The first attribute of the new states simulates the sequence of states for a given string in the DFA A. The second attribute simulates all possible computation that we could take in the DFA A but in reverse (starting by the final states). We start from (q_0, F), because if we read ε we stay at q_0 and we can reach F from F through ε. The final states are $F' = \{(q, S) \in Q' \mid q \in S\}$, because it means that we have reached q by reading some string w and we can reach one of the final states through reading a string of twice the length of w. Therefore, $\frac{1}{2}(L)$ is regular.