University of Illinois at Urbana-Champaign
Department of Computer Science

Midterm 1 Solutions

CS 373: Theory of Computation
Spring 2011

Name:	
Netid:	

- Print your name and netid, *neatly* in the space provided above; print your name at the upper right corner of *every* page. Please print legibly.
- This is a *closed book* exam. No notes, books, dictionaries, calculators, or laptops are permitted.
- You are free to cite and use any theorems from class or homeworks without having to prove them again.
- Write your answers in the space provided for the corresponding problem. Let us know if you need more paper.
- Suggestions: Read through the entire exam first before starting work. Do not spend too much time on any single problem. If you get stuck, move on to something else and come back later.
- If you run short on time, remember that partial credit will be given.
- If any question is unclear, ask us for clarification.

Question	Points	Score
Problem 1	25	
Problem 2	15	
Problem 3	15	
Problem 4	15	
Problem 5	15	
Problem 6	15	
Total	100	

1. Short Problems (25 points)

Give answers to each of the following questions, including a short justification where requested.

(a) Give a simple description of the strings belonging to the language described by the regular expression $0^* \cup (0^*10^*10^*10^*)^*$. (4 points)

(b) If A and B are disjoint, nonregular languages ($A \cap B = \emptyset$), then must $A \cup B$ be nonregular? Briefly justify your answer. (4 points)

(c) Let $d(w)$ denote $|n - m|$ in which n is the number of 0s in w and m is the number of 1s in w. Is $A = \{w \in \{0,1\}^* \mid d(w) = 0 \bmod 3\}$ regular? Briefly justify your answer. (4 points)

(d) Give a regular expression that describes all strings w for which $|w|$ is even or w contains the substring 10101. (4 points)

(e) If A is a regular language, then is $A^C = \Sigma^* \setminus A$ a regular language? Show your reasoning. (4 points)

(f) Is the set of all NFAs over an alphabet $\Sigma = \{0,1\}$ finite, countably infinite, or uncountably infinite? (3 points)

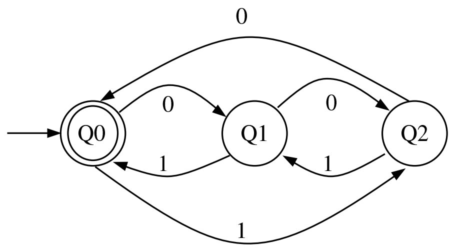
(g) Is the set of all languages over an alphabet $\Sigma = \{0,1\}$ finite, countably infinite, or uncountably infinite? (2 points)

Solutions:

(a) A = $\{w \in \{0,1\}^* \mid$ number of 1s in w is a multiple of 3.$\}$

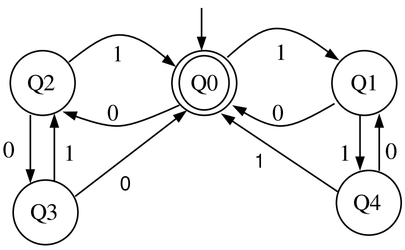
(b) False. Consider $A = \{0^m1^n \mid m < n\}$ and $B = \{0^p1^q \mid p \geq q\}$.
A and B are non-regular languages. However, A∪B = 0^*1^*, which is regular.

(c) Yes. A DFA that recognizes A can be drawn as,



where, Q0 - start and accept state (same number of 0s and 1s)
Q1 - |x-y| mod 3 is 1.
Q2 - |x-y| mod 3 is 2.
x - number of 1s
y - number of 0s

OR



where, Q0 - start and accept state (same number of 0s and 1s)
Q1 - (number of 1s) mod 3 is 1.
Q4 - (number of 1s) mod 3 is 2.
Q2 - (number of 0s) mod 3 is 1.
Q3 - (number of 0s) mod 3 is 2.

(d) $(\Sigma\Sigma)^* \cup (\Sigma^* 10101 \Sigma^*)$

(e) Yes. Let M = (Q,Σ, δ, q_0,F) be a DFA that recognizes A. Then a DFA that accepts the complement of A, i.e., $\Sigma^* - A$ can be obtained by swapping its accepting states with its non-accepting states. $M^c = (Q, \Sigma, \delta, q_0, Q/F)$ is a DFA that accepts $\Sigma^* - A$.

(f) Countably infinite.They can be written as a finite-length string of symbols over some alphabet.

(g) Uncountably infinite. We show this by constructing a bijective mapping $f : L \to B$. For each language $A \in L$, we can construct a unique element in B called the characteristic sequence. Let $\Sigma^*_{0,1} = s_1, s_2, s_3, ...$, then the i^{th} bit of the characteristic sequence of A is 1 if $s_i \in A$, and 0 if $s_i \notin A$. Note,

- The empty language has the characteristic sequence 000000 . . .
- The language $\Sigma^*_{0,1}$ has the characteristic sequence 1111 . . .

The mapping f is bijective in that any possible language in L has a unique sequence in B and any sequence in B uniquely denes a language in L.

2. Preserving Regularity (15 points)

For any language A, let

$$FLIP(A) = \{w \in \Sigma^* \mid \text{there exist } x, y \in \Sigma^* \text{ such that } xy \in A \text{ and } w = yx\}.$$

For example, if $A = \{01, 011\}$, then $FLIP(A) = \{01, 10, 011, 110, 101\}$.

Prove that if A is a regular language, then $FLIP(A)$ is also a regular language.

[Hint: Remember the power of nondeterminism.]

Solution:

Let $(Q_A, \Sigma_A, \delta_A, q_{0,A}, F_A)$ be a DFA that recognizes the regular language A. We will construct an NFA $(Q_F, \Sigma_F, \delta_F, q_{0,F}, F_F)$ that recognizes $FLIP(A)$.

- $Q_F = (Q_A \times Q_A \times \{0, 1\}) \cup q_{0,F}$, the first copy of Q_A stores the guess that we made about where the string y started, the second copy of Q_A stores the current state, and the third member of the ordered pair is 0 if the machine believes it is still reading y, and 1 if it believes that x has started. The state $q_{0,F}$ is a new start state.
- $\Sigma_F = \Sigma_A$
- $\delta_F(q_{0,F}, \varepsilon) = \{(q, q, 0) \mid q \in Q_A\}$, we guess that the machine could be in any state at the start of y.
- $\delta_F((q_1, q_2, 0), a) = \{(q_1, \delta_A(q_2, a), 0)\}$, we can never be sure that y has ended, so whenever a character is read, one of our guesses must be that we are still reading y, so the second member of the ordered pair is transitioned.
- $\delta_F((q_1, q_2, 1), a) = \{(q_1, \delta_A(q_2, a), 1)\}$, if we have already made a guess, then we just continue transitioning the second member of the ordered pair.
- $\delta_F((q_1, q_2 \in F_A, 0), \varepsilon) = \{(q_1, q_{0,F}, 1)\}$, if q_2 is a final state of the original DFA, then we may have hit the end of y, so we must non-deterministically guess that x has started.
- $q_{0,F} = q_{0,F}$
- $F_F = \{(q, q, 1) \mid q \in Q_A\}$, since x must end where y began, the first two members of the ordered pair must be the same.

3. Induction (15 points)

Prove by induction: If A is any language that contains exactly n strings, then A is regular. Express your arguments in terms of appropriate sets and functions.

Solution:

There are two base cases. The first is $n = 0$, the language that has no strings. Define a DFA $M_0 = (Q, \Sigma, \delta, q_0, F)$ where $F = \emptyset$. Since $F = \emptyset$, the machine never accepts, so $|L(M_0)| = 0$, meaning that any language that contains no strings is regular.

The other base case is $n = 1$. Let M_1 be a language containing exactly one string in Σ^*. Let $w = a_1 a_2 \ldots a_m$ be the only member of M_1. Define an NFA $M_1 = (Q_1, \Sigma, \delta_1, q_{0,1}, F_1)$ where

* $Q_1 = \{q_0, q_1, \ldots q_m\}$, $m + 1$ states denoting progress along the string w
* $\delta_1(q_i, a) = \{q_{i+1}\}$ if $a = a_{i+1}$,
* $q_{0,1} = q_0$
* $F_1 = \{q_m\}$

This accepts only the string where the ith character is a_i, and only w meets this criterion, so $L(M_1) = w$, so any language with exactly one string is regular.

For an inductive hypothesis, assume that any language with less than n strings is regular.

For the inductive step, let L_n be a language that contains exactly n strings. Let w be a string in L_n. The language $L_{n-1} = L_n \setminus \{w\}$ is regular by the inductive hypothesis, so there exists an NFA $M_{n-1} = (Q_{n-1}, \Sigma, \delta_{n-1}, q_{0,n-1}, F_{n-1})$ such that $L(M_{n-1}) = L_{n-1}$. By the base case, there exists an NFA $M_1 = (Q_1, \Sigma, \delta_1, q_{0,1}, F_1)$ such that $L(M_1) = \{w\}$.

We can now define a new NFA $M_n = (Q_n, \Sigma, \delta_n, q_{0,n}, F_n)$.

* $Q_n = Q_{n-1} \cup Q_1 \cup \{q_{0,n}\}$
* $\delta_n(q \in Q_1, a) = \delta_1(q, a)$
* $\delta_n(q \in Q_{n-1}, a) = \delta_{n-1}(q_a)$
* $\delta_n(q_{0,n}, \varepsilon) = \{q_{0,1}, q_{0,n-1}\}$
* $q_{0,n} = q_{0,n}$
* $F_n = F_{n-1} \cup F_1$

This machine accepts any strings in $L_{n-1} \cup L_1 = L_n$, as M_{n-1} and M_1 are submachines in M_n, and the string is processed by both M_{n-1} and M_1 simultaneously. Since M_n recognizes L_n, L_n is regular.

Therefore, any language with exactly n strings is regular.

4. Pumping Lemma (15 points)

Use the pumping lemma to prove that

$$\{(ba)^n b^n \in \{a,b\}^* \mid n \geq 0\}$$

is nonregular.

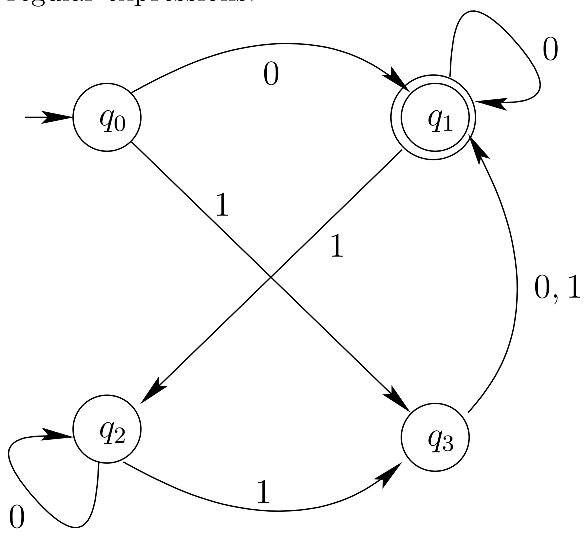
Solution:

Let A be the language in the problem and assume that it's regular. Let p be the pumping length from the pumping lemma.

Consider $w = (ba)^{2p}b^{2p} \in A$, and note $|w| > p$. We can break w into xyz such that $y \neq \varepsilon$ and $|xy| \leq p$. xy must be a prefix of $(ba)^{p/2}$ in w_1, so y could contain an a or a b. If y contains an a, note that xy^0z has fewer than $2p$ as, but there are $2p$ bs at the end of w, which is unchanged. Hence, $xy^0z \notin A$. Else $y = b$, but then xy^0z contains the string of two consecutive as, so $xy^0z \notin A$. Thus no matter how we partition w into xyz, we cannot pump it. We have a contradiction and A cannot be regular.

5. DFA to Regular Expression (15 points)

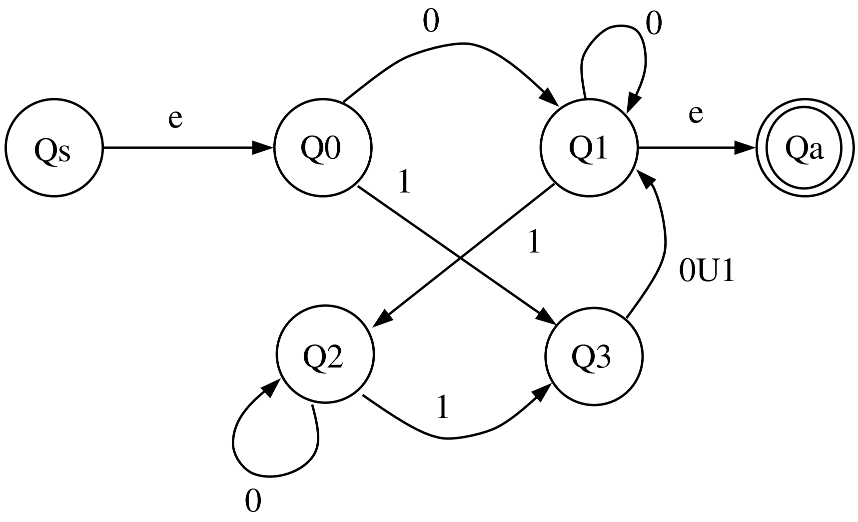
   Convert the following DFA into a regular expression that describes the same language. Use the method of GNFAs. Show the complete, resulting GNFA after each deletion. There is no need to simplify the regular expressions.



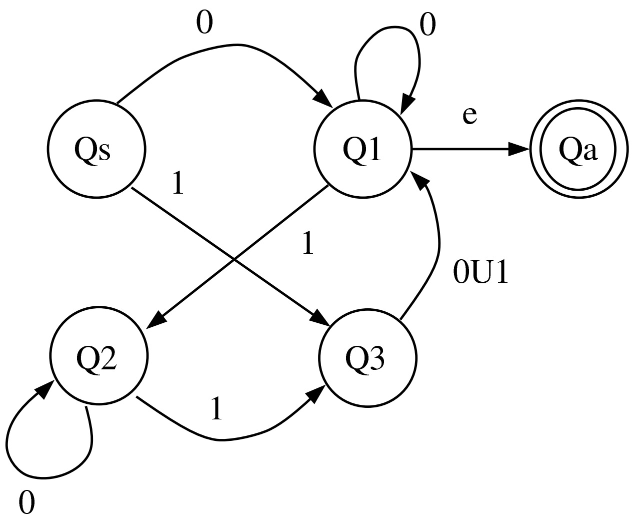
Solution:

Edges that correspond to transitions on $\emptyset$ are not drawn here. If no edge is drawn between two states in a GNFA diagram, then the edge between them corresponds to the empty set.

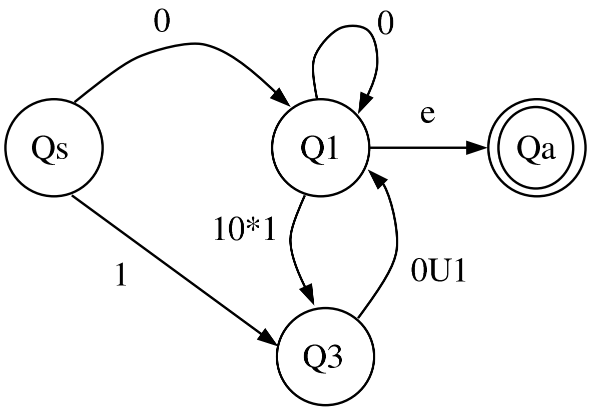
1. Convert DFA to GNFA



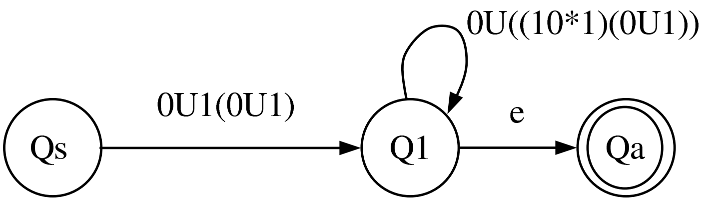
2. Removing Q0



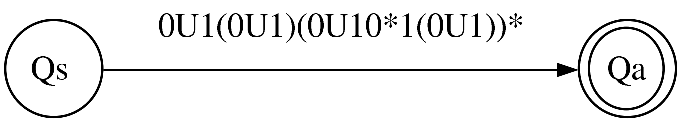
3. Removing Q2



4. Removing Q3



5. Removing Q1



6. Prove or Disprove (15 points)

   Let A denote a language. If there exists some $p \in \mathbb{N}$ such that for all $s \in A$ for which $|s| \geq p$, there exists strings x, y, and z such that $s = xyz$ and $xy^iz \in A$ for all integers $i \geq 0$, then A is a regular language.

   Solution 1:

   The claim is false. Consider $A = \{0^n1^n \mid n \geq 0\}$, which we know to be nonregular. This language fulfills the conditions above with $p = 2$. Consider any string $s = 0^k1^k \in A$, $k \geq 1$. Let $x = s$ and $y = z = \varepsilon$. Then $xy^iz = x = s \in A$ for all i. Thus the implication does not hold.

   Solution 2:

   This is false. In section, we showed that the language $A = \{a^ib^jc^k \in a^*b^*c^* \mid i, j, k \geq 0, i = 1 \to j = k\}$ fulfills the conditions of the pumping lemma, but is nonregular. The pumping lemma is stronger than the conditions above, so this language also disproves our claim.

   We will show that A fulfills the conditions of the pumping lemma. Let the pumping length $p = 2$, and take $s \in A$, $|s| \geq p$. If $s \in b^*c^*$, $s \in ab^*c^*$, or $s \in aaaa^*b^*c^*$, let x be empty and y be the first character of s. If $s \in aab^*c^*$, then let x be empty and $y = aa$. Then $xy^iz \in A$ for all i in all cases.

   Now we will show that A is nonregular. Assume that A is regular. Consider a homomorphism $h : \{a, b, c\} \to \{0, 1\}$ such that $h(a) = \varepsilon$, $h(b) = 0$, and $h(c) = 1$. Notice that we can create the nonregular language $h(A \cap ab^*c^*) = \{0^n1^n \mid n \geq 0\}$ by closure properties. This is contradiction, so A must not be regular.