CS 341, Spring 2012
Solutions for Midterm 1

1. (a) True. See slide 2-111.
   (b) True. HW 4, problem 5(a).
   (c) True. HW 4, problem 5(c).
   (d) False. 1^* is regular since it has a regular expression, but this language is infinite.
   (e) True. By HW 2, problem 3, we know $\overline{A}$ is regular. Since $\overline{A}$ and B are regular, then $\overline{A} \cup B$ is regular by Theorem 1.25. Theorem 1.49 then implies $(\overline{A} \cup B)^*$ is regular.
   (f) False. See HW 6, problem 2(a).
   (g) False. For example, $A = \{0^n 1^n 0^n \mid n \geq 0\}$ is a subset of $B = L((0 \cup 1)^*)$, but A is non-context-free and B is context-free.
   (h) True. Use the pumping lemma with string $a^r b^r$, where $r = \max(p, 3)$ and p is the pumping length.
   (i) True. Since A has a regular expression, A is a regular language by Theorem 1.54. Then Corollary 2.32 implies A is also context-free, so it has a CFG. Theorem 2.9 then ensures that A has a CFG in Chomsky normal form.
   (j) True. Suppose A is non-context-free but regular. But then Corollary 2.32 implies A is context-free, which is a contradiction.

2. (a) $b^*ab^* \cup b^*ab^*(a \cup \varepsilon)b^*$. Another regular expression is $b^*(a \cup ab^*(a \cup \varepsilon))b^*$. There are infinitely many regular expressions for the language.
   (b) $G' = (V', \Sigma, R', S_0)$, where $V' = V \cup \{S_0\}$, S_0 is the (new) starting variable, Σ is the same alphabet of terminals as in G, and $R' = R \cup \{S_0 \to SS_0 \mid \varepsilon\}$.
   (c) $M_3 = (Q_3, \Sigma, \delta_3, q_3, F_3)$, where $Q_3 = Q_1 \times Q_2$; Σ is the same alphabet as M_1 and M_2 have; the transition function δ_2 satisfies $\delta((q, r), \ell) = (\delta_1(q, \ell), \delta_2(r, \ell))$ for $(q, r) \in Q_3$ and $\ell \in \Sigma$; the starting state $q_3 = (q_1, q_2)$; and $F_3 = (Q_1 \times F_2) \cap (F_1 \times Q_2)$, which also can be written as $F_1 \times F_2$.
   (d) After one step, the CFG is then

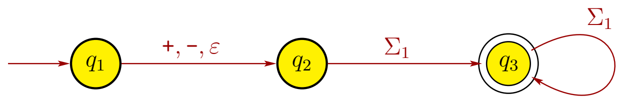
$$\begin{aligned} S_0 &\to S \\ S &\to A1SA \mid 1SA \mid A1S \mid 1S \mid A0 \mid 0 \mid \varepsilon \\ A &\to 0S0 \end{aligned}$$

3. (a) A regular expression for L_1 is

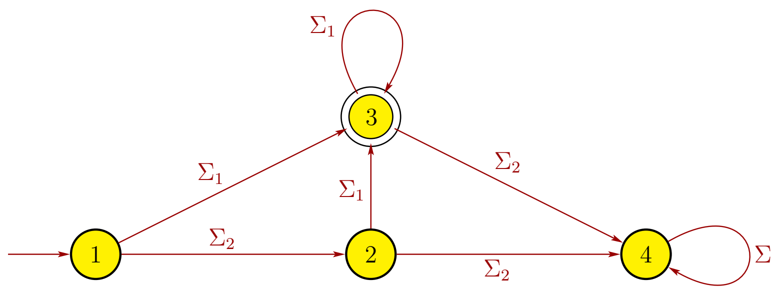
$$R_1 = (\,\texttt{+} \cup \texttt{-} \cup \varepsilon\,)\,\Sigma_1\,\Sigma_1^*$$

   where $\Sigma_1 = \{\,\texttt{0},\,\texttt{1},\,\texttt{2},\,\ldots,\,\texttt{9}\,\}$ as defined in the problem.

(b) An NFA for L_1 is



(c) Define $\Sigma_2 = \{\,\texttt{-},\,\texttt{+}\,\}$, as given in the problem. Then a DFA for L_1 is

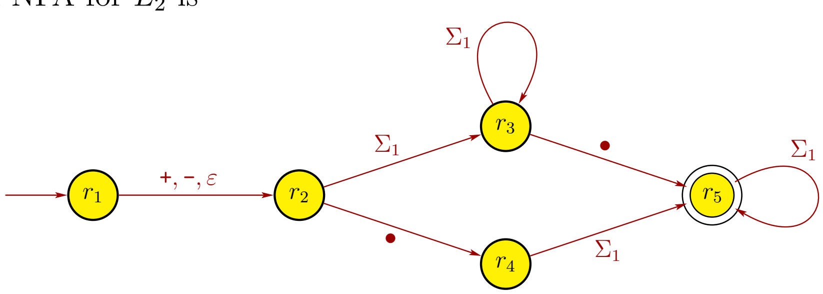


(d) A regular expression for L_2 is

$$R_2 = (\,\texttt{+} \cup \texttt{-} \cup \varepsilon\,)(\Sigma_1\,\Sigma_1^*\,\texttt{.}\,\Sigma_1^* \;\cup\; \texttt{.}\,\Sigma_1\,\Sigma_1^*)$$

Note that the regular expression $(\,\texttt{+} \cup \texttt{-} \cup \varepsilon\,)\,\Sigma_1^*\,\texttt{.}\,\Sigma_1^*$ is not correct since it can generate the strings "`.`", "`+.`" and "`-.`", which are not valid floating-point numbers.

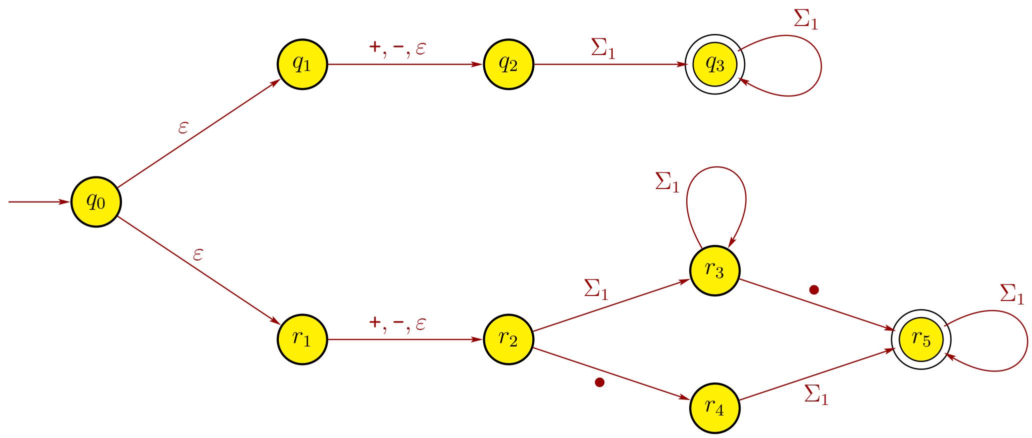
(e) An NFA for L_2 is



(f) Note that $L = L_1 \cup L_2$, so a regular expression for L is

$$R_3 = R_1 \cup R_2$$

(g) We can construct an NFA for L by taking the union of the NFA's for L_1 and L_2 as follows:

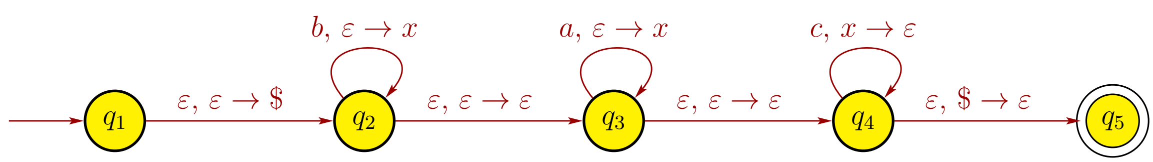


There are many other correct answers for this and the other parts.

4. (a) $G = (V, \Sigma, R, S)$ with set of variables $V = \{S, X\}$, where S is the start variable; set of terminals $\Sigma = \{a, b, c\}$; and rules

$$\begin{aligned} S &\rightarrow bSc \mid X \\ X &\rightarrow aXc \mid \varepsilon \end{aligned}$$

(b) PDA



For every b and a read in the first part of the string, the PDA pushes an x onto the stack. Then it must read a c for each x popped off the stack.

5. Language A is nonregular. We prove this by contradiction. Suppose that A is a regular language. Let p be the "pumping length" of the Pumping Lemma. Consider the string $s = a^p b a^p b$. Note that $s \in A$, and $|s| = 2p + 2 > p$, so the Pumping Lemma will hold. Thus, there exists strings x, y, and z such that $s = xyz$ and

(a) $xy^iz \in A$ for each $i \geq 0$,

(b) $|y| > 0$,

(c) $|xy| \leq p$.

Since the first p symbols of s are all a's, the third property implies that x and y consist of only a's. So z will be the rest of the a's from the beginning, followed by ba^pb. The second property states that $|y| > 0$, so y has at least one a. More precisely, we can then say that

$$x = a^j \text{ for some } j \geq 0,$$

$$\begin{aligned} y &= a^k \text{ for some } k \geq 1, \\ z &= a^m b a^p b \text{ for some } m \geq 0. \end{aligned}$$

Since $a^p b a^p b = s = xyz = a^j a^k a^m b a^p b = a^{j+k+m} b a^p b$, we must have that

$$j + k + m = p.$$

The first property implies that $xy^2z \in A$, and

$$\begin{aligned} xy^2z &= a^j a^k a^k a^m b a^p b \\ &= a^{p+k} b a^p b \end{aligned}$$

since $j+k+m = p$. If xy^2z has odd length, then it cannot be split into equal halves, so $xy^2z \notin A$. If xy^2z has even length, then the first half contains only a's since $k \geq 1$, but the second half contains 2 b's. Hence, $xy^2z \notin A$, so we get a contradiction. Therefore, A is a nonregular language.