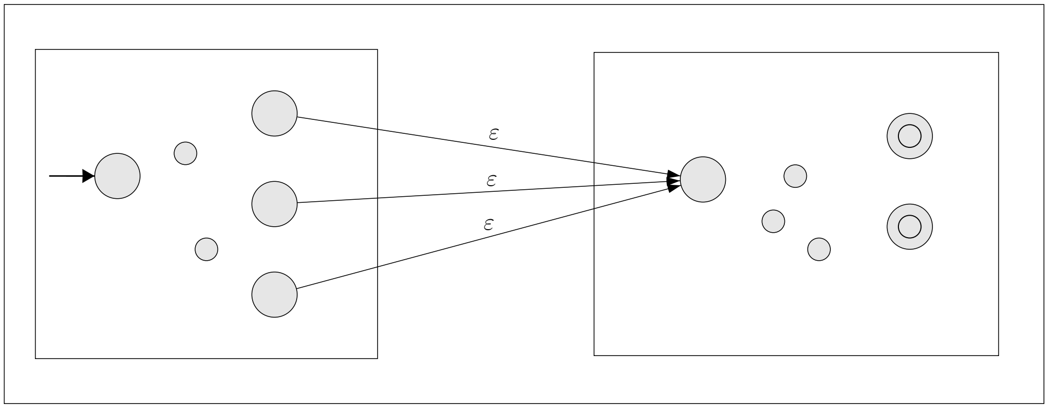
CS 341-452, Spring 2019
Solutions for Midterm, eLearning Section

1. (a) True. By Theorem 2.9. The fact that A is infinite is irrelevant.
   (b) True. Because B is finite, we have that $A \cap B$ is also finite, so it is regular by slide 1-95.
   (c) False. A TM can loop on w.
   (d) False. For example, let $A = \{abc\}$ and $B = \{a^n b^n c^n \mid n \geq 0\}$, so $A \subseteq B$. Because A is finite, it is regular (slide 1-95), so it is also context-free by Corollary 2.32. But B is not context-free by slide 2-96.
   (e) True. By Homework 5, problem 3(b).
   (f) False. $A = \{ a^n b^n c^n \mid n \geq 0 \}$ is nonregular and not context-free.
   (g) False. For example, let $A = \{ 0^n 1^n \mid n \geq 0 \}$ is context-free (see slide 2-5) and infinite.
   (h) True. Because A is finite, it is regular by the theorem on slide 1-95 of the notes. Corollary 2.32 then ensures that A is regular, so $\overline{A}$ is also regular (Homework 2, problem 3). Corollary 2.32 implies $\overline{A}$ is context-free.
   (i) False. For example, let $A = \{a^n b^n c^n \mid n \geq 0\}$, and let $B = \Sigma^*$ for $\Sigma = \{a, b, c\}$. Thus, we have that $A \subseteq B$. Because B has a regular expression (e.g., $(a \cup b \cup c)^*$), B is regular by Kleene's theorem. But A is not context-free (slide 2-96).
   (j) False. $a^* b^* a^*$ generates the string $abbaaa \notin \{ a^n b^n a^n \mid n \geq 0 \}$. In fact, the language $\{ a^n b^n a^n \mid n \geq 0 \}$ is not regular, so it does not have a regular expression.

2. (a) A regular expression is $a^*ba^*ba^*b(a \cup b)^* \cup b^*ab^*ab^*$. There are infinitely many other correct regular expressions.
   (b) The rules that violate Chomsky normal form are
      - $S \to ba$ because a rule cannot go to 2 terminals;
      - $X \to YS$ because the starting variable S cannot be on the right side of a rule;
      - $X \to \varepsilon$ because it is an ε-rule;
      - $Y \to X$ because it is a unit rule;
      - $Y \to aX$ because the right side has a mix of terminals and variables.

   (c)
   (d) As given on slide 1-63, $A_1 \circ A_2$ has the following NFA N:

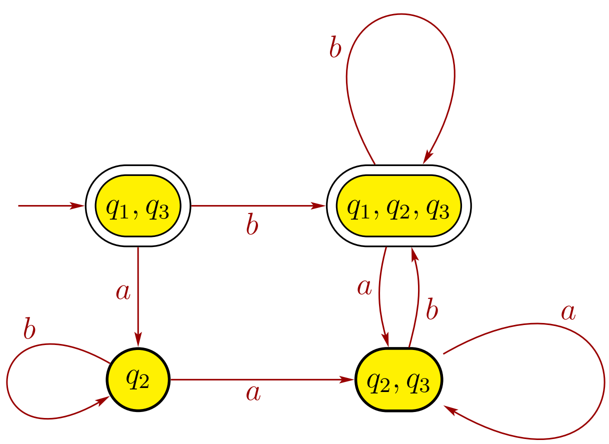
N



(e) Homework 5, problem 3a. We are given a CFG $G_1 = (V_1, S, R_1, S_1)$ for language A_1, and a CFG $G_2 = (V_2, S, R_2, S_2)$ for language A_2. We can then define a CFG $G_3 = (V_3, S, R_3, S_3)$ for $A_1 \cup A_2$ with $V_3 = V_1 \cup V_2 \cup \{S_3\}$, where $S_3 \notin V_1 \cup V_2$, and $R_3 = R_1 \cup R_2 \cup \{S_3 \to S_1, S_3 \to S_2\}$.

3. $q_1 10\#1100 \quad xq_3 0\#1100 \quad x0q_3\#1100 \quad x0\#q_5 1100 \quad x0q_6\#x100 \quad xq_7 0\#x100 \quad q_7 x0\#x100$ $xq_1 0\#x100 \quad xxq_2\#x100 \quad xx\#q_4 x100 \quad xx\#xq_4 100 \quad xx\#x1q_{\text{reject}}00$

4. DFA

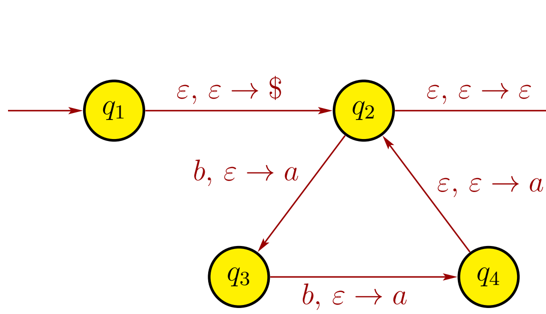


5. (a) $G = (V, \Sigma, R, S)$, with $V = \{S, X\}$, $\Sigma = \{a, b, c\}$, start variable S and rules

$$\begin{aligned} S &\to bbSaaa \mid X \\ X &\to cX \mid \varepsilon \end{aligned}$$

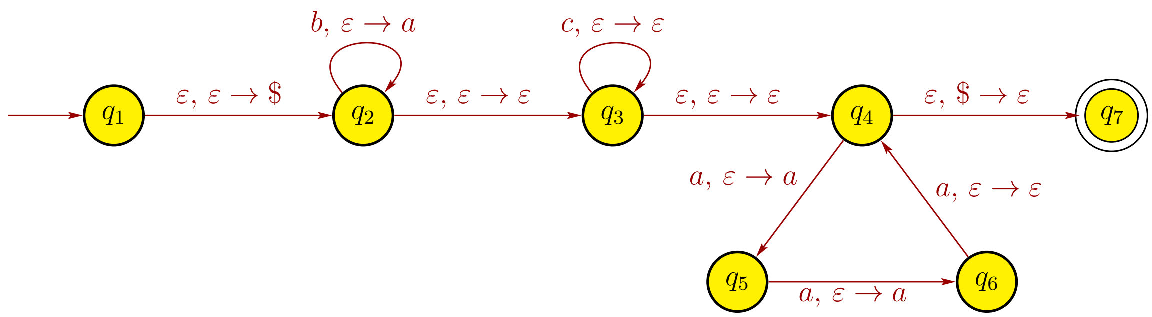
There are infinitely many other correct CFGs for A.

(b) PDA



The loop from $q_2 \to q_3 \to q_4 \to q_2$ reads a b on the first two transitions, but reads ε on the third transition; all three transitions push an a. This has the effect of pushing 3 a's for every 2 b's that are read. The loop on q_5 reads c's, but doesn't alter the stack because we don't have to match the c's with anything. Next, the loop on q_6 just reads an a to match every a on the stack. Finally, the transition from q_6 to q_7 makes sure there aren't any leftover a's in the stack.

Another PDA for the language is as follows:



For the second PDA, the loop on q_2 pushes an a on the stack for each b read. The loop on q_3 reads c's, but doesn't alter the stack because we don't have to match the c's with anything. Finally, the loop $q_4 \to q_5 \to q_6 \to q_4$ reads an a on each of the three transition, but only pops an a on each of the first two transitions. Because an a was pushed onto the stack for every b read, the loop $q_4 \to q_5 \to q_6 \to q_4$ ensures that three a's are read for every two b's, as required. Finally, the transition from q_4 to q_7 makes sure there aren't any leftover a's in the stack.

There are infinitely many other correct PDAs for A.

6. This is Homework 2, problem 4. We prove this by contradiction. Suppose that $\overline{M}$ is not a minimal DFA for $\overline{A}$. Then there exists another DFA D for $\overline{A}$ such that D has strictly fewer states than $\overline{M}$. Now create another DFA D' by swapping the accepting and non-accepting states of D. Then D' recognizes the complement of $\overline{A}$. But the complement of $\overline{A}$ is just A, so D' recognizes A. Note that D' has the same number of states as D, and $\overline{M}$ has the same number of states as M. Thus, because we assumed that D has strictly fewer states than $\overline{M}$, then D' has strictly fewer states than M. But since D' recognizes A, this contradicts our assumption that M is a minimal DFA for A. Therefore, $\overline{M}$ is a minimal DFA for $\overline{A}$.

7. The language $A = \{ b^{2n} c^k a^{3n} \mid n \geq 0, k \geq 0 \}$ is not regular. To prove this, suppose that A is a regular language. Let p be the pumping length, and consider the string $s = b^{2p} a^{3p} \in A$. Note that $|s| = 5p \geq p$, so the pumping lemma implies we can write $s = xyz$ with $xy^iz \in A$ for all $i \geq 0$, $|y| > 0$, and $|xy| \leq p$. Now, $|xy| \leq p$ implies that x and y have only b's (together up to p in total) and z has the rest of the b's at the beginning, followed by a^{3p}. Hence, we can write

$$\begin{aligned} x &= b^j, \text{ for some } j \geq 0, \\ y &= b^\ell, \text{ for some } \ell \geq 0, \\ z &= b^m b^p a^{3p}, \text{ for some } m \geq 0. \end{aligned}$$

Because $xyz = b^j b^\ell b^m b^p a^{3p} = s = b^{2p} a^{3p}$, we have that $j + \ell + m + p = 2p$, or equivalently, $j + \ell + m = p$. Also, $|y| > 0$ implies $\ell > 0$. Now consider the string $xyyz = b^j b^\ell b^\ell b^m b^p a^{3p} = b^{2p+\ell} a^{3p}$ because $j + \ell + m = p$. Note that $xyyz \notin A$ because the numbers of b's and a's don't have the right relationship because $\ell > 0$, which contradicts (i). Hence, A is not a regular language.