Closed Book, Closed Notes Time Limit: 1 hour 50 minutes

1. Answer True or False to the following questions and briefly JUSTIFY each answer.

   (a) If a PDA M actually pushes a symbol on its stack then $L(M)$ is not regular.

   (b) For every infinite regular language L over Σ there are strings $x, y, z \in \Sigma^*$ such that $|y| \geq 1$ and for every $k \geq 0$, $xy^kz \in L$.

   (c) There is no algorithm to tell, given an arbitrary string P whether or not P is the ASCII for a syntactically correct C program.

   (d) If L is a CFL and $L = K \cap R$ for a regular language R then K is a CFL.

   (e) If L is not a CFL and $L = K \cap R$ for a regular language R then K is not a CFL.

   (f) There is no algorithm to tell, given an arbitrary program P and an input x, whether or not P runs forever on input x.

   (g) If the stack in a PDA M only has capacity for 100 characters on any input then $L(M)$ is regular.

   (h) If L is accepted by some PDA M then L^R is accepted by some PDA M'.

   (i) There is no algorithm to tell, given an arbitrary CFG G and input x whether or not $x \in L(G)$.

   (j) Context-free languages are more interesting than regular languages.

2. Classify each of the following sets as:

   A Both Regular and Context-Free,

   B Context-Free, but not regular,

   C Neither context-free nor regular but membership in the language can be decided by an algorithm.

   D Undecidable

   You do not need to justify your answers.

   (a) $\{a^nb^m \mid m = 2n + 1\}$ ..... A B C D

   (b) $\{a^na^m \mid m = 2n + 1\}$ ..... A B C D

   (c) $\{a^nb^m \mid m \equiv n \pmod 3\}$ ..... A B C D

   (d) $\{xcx^R \in \{a, b\}^* \mid x$ has an even number of a's$\}$ ..... A B C D

   (e) $\{a^ib^jc^k \mid k = i + j\}$ ..... A B C D

   (f) $\{a^ib^ja^k \mid k \neq i$ or $k \neq j\}$ ..... A B C D

   (g) $\{a^ib^ja^k \mid k \neq i$ and $k \neq j\}$. ..... A B C D

   (h) $\{a^nb^na^mb^m \mid m, n \geq 0\}$. ..... A B C D

   (i) $\{a^nb^ma^nb^m \mid m, n \geq 0\}$. ..... A B C D

   (j) $\{xy \in \{a, b\}^* \mid x \neq y\}$. ..... A B C D

3. Let $L = \{0^n 1^m \mid m \text{ is an integer multiple of } n\}$.
   Use the Myhill-Nerode theorem to show that L is not regular.

4. Let $G = (V, \Sigma, R, S)$ be the context-free grammar with the following set of rules:

$$S \to aSaSb \mid aSbSa \mid bSaSa \mid SS \mid e$$

   (a) Use one of the general constructions that convert a CFG to a PDA to give a PDA M such that $L(M) = L(G)$.

   (b) Show each step of a computation of your PDA (list the stack contents, state, and amount of input remaining) that accepts input $aabbaaaab$.

   (c) (Bonus) What language does this grammar generate?

5. Let $\#_a(x)$ be the number of a's in x and $\#_b(x)$ be the number of b's in x.
   Let $L = \{x \in \{a, b\}^* \mid \#_a(x) = \#_b(x) \text{ and both are even}\}$.

   (a) Give the formal definition of a context-free grammar G such that $L(G) = L$.

   (b) Give a parse tree for the string: $aabbbaab$

   (c) Write out the leftmost derivation corresponding to the parse tree in part (b).

6. Let $L = \{a^m b^n c^p \mid 0 \leq m < n < p\}$. Prove that L is not a context-free language.

7. For any context-free grammar $G = (V, \Sigma, R, S)$, we say that a nonterminal $A \in V$ is *useful* if there is a derivation $S \Rightarrow^* xAy \Rightarrow^* w$ for $w \in \Sigma^*$ and $x, y \in V^*$; otherwise it is *useless*. Suppose that G has the following rules:

$$\begin{aligned}
S &\to AC \mid BS \mid B \\
A &\to aA \mid aT \\
B &\to CF \mid b \\
C &\to cC \mid D \\
D &\to aD \mid BD \mid C \\
E &\to aA \mid BSA \\
T &\to bB \mid b \\
U &\to bA \mid a \mid Wb \\
W &\to Ub \mid a \mid Bb
\end{aligned}$$

   (a) Which non-terminals of G are useful?

   (b) Modify G to get an equivalent grammar G' whose nonterminals consist of precisely the nonterminals of G that are useful. (Don't remove any other non-terminals.)

   (c) Describe a reasonably efficient algorithm that will remove all useless non-terminals from a grammar and thus show that any grammar is equivalent to one with only useful non-terminals.