Problem Set 3

Fall 09

Due: Thursday Oct 22 at 11:00 AM in class (i.e., Room 103 Talbot Lab)
Please follow the homework format guidelines posted on the class web page:
`http://www.cs.uiuc.edu/class/fa09/cs373/`

1. CFG construction. [**Points**: 30]
Construct CFGs for the following languages. Give a brief explanation of how your grammar works and what each nonterminal stands for.

$$A = \{a^i b^j \mid i \neq j\}$$
$$B = \{a^i b^j c^k \mid i \leq j \text{ or } j \leq k\}$$
$$C = \{x \in \{0,1\}^* \mid x \text{ is not of the form } ww\}$$
$$D = \{x \in \{a,b\}^* \mid x \text{ contains twice as many } a\text{'s than } b\text{'s}\}$$
$$E = \{x \in \{a,b\}^* \mid x \text{ is not a palindrome and } |x| \text{ is even}\}$$
$$F = \{a^i b^j \mid 2i+3 \leq j \leq 4i+5\}$$

2. CFG interpretation. [**Points**: 15]
What is the language of the following CFG? Justify your answer.

$$S \to A1B$$
$$A \to 0A \mid \epsilon$$
$$B \to 0B \mid 1B \mid \epsilon$$

3. Proof. [**Points**: 30]
Consider the context-free grammar:

$$S \to 0S1S \mid 1S0S \mid \epsilon$$

Describe $L(G)$ in English, and prove that your answer is correct, i.e. G generates all strings you've described and only those strings.

4. PDA stack reduction. [**Points**: 40]

   (a) Construct a PDA with input alphabet $\Sigma = \{a,b,c,d\}$ and a stack alphabet $\Gamma = \{A,B,C\}$ that accepts the language $\{wdw^r \mid w \in \{a,b,c\}^*\}$. Explain how your PDA works.

   (b) Construct a PDA with input alphabet $\Sigma = \{a,b,c,d\}$ and a stack alphabet $\Gamma = \{E,F\}$ that accepts the language $\{wdw^r \mid w \in \{a,b,c\}^*\}$. Explain how your PDA works.

   (c) Prove that if P is a PDA, then there exists another PDA, say P', with only two stack symbols such that $L(P') = L(P)$.

5. Chomsky Normal Form. [**Points**: 30]
Let L be the language generated by the grammar G below:

$$S \to XY \mid YYY$$
$$X \to Yb \mid \epsilon$$
$$Y \to aY \mid X$$

(a) Eliminate all ϵ-productions from G (obtaining a grammar G_1 for $L - \{\epsilon\}$).

(b) Eliminate all unit productions from G_1, obtaining G_2.

(c) Put G_2 into Chomsky Normal form, obtaining G_3.

6. Index of a derivation. [**Points**: 30]
A *derivation* is a sequence of substitutions that generate a string from some non-terminal of G, for example $S \implies SaSB \implies bbaSB \implies bbaaB \implies bbaaa$ could be a derivation for string $bbaaa$ from non-terminal S in some grammar. Note that in each step of derivation we just substitute <u>one</u> non-terminal, for example in the second step of derivations above, the first S has been substituted by bb. The outcome of each step in a derivation, is called a *sentential form* of that derivation, for example $SaSB$, $bbaSB$, $bbaaB$, and $bbaaa$ are four sentential forms of the derivation above.

The *index* of a derivation is the maximum number of nonterminals in any sentential form of that derivation. The index of a string w generated by grammar G, denoted by $idx_G(w)$, is the smallest index of any derivation that derives w from the start symbol of G. The index $idx(G)$ of a grammar G is defined as $idx(G) = \max_w\{idx_G(w) \mid w \in L(G)\}$.

Show that the index of G over $\{a, b\}$ with productions $S \to aSb \mid SS \mid \epsilon$ is infinite.

7. Non-CFL. [**Points**: 20]
Prove $A = \{0^j1^k \mid k > j^2\}$ is not context-free.

8. PDA construction. [**Points**: 30]
Given $A = \{(0^i1^j)^k \mid 1 \leq i < j,\ k < 2i\}$, construct a PDA for the complement of A.

9. Regularity. [**Points**: 30]
Given a CFG $X = (V, \Sigma, R, S)$ such that for all $v \in V$, there is a single $r \in R$ with v on the left side, show that $L(X)$ is accepted by some DFA.