CS3133

HW#5

DUE: Friday, September 29

1. (4 points) Rewrite the following CFG in Chomsky Normal Form.

$$S \to 0AS \mid 0$$
$$A \to S1A \mid SSSS \mid 10$$

2. (8 points) Prove or give a counterexample to the following.

CONJECTURE: For any context free language L, if $\varepsilon \notin L$ then there is a context free grammar G such that $L = L(G)$ and for any $w \in L$, there is a derivation $S \xrightarrow[G]{*} w$ of exactly $(2*|w|)-1$ steps.

3. (8 points) Is the language $\{0^n10^n10^{2n} \mid n \geq 0\}$ context free? Justify your answer.

CS3133

Solutions for HW#5

1. In the first pass we assure that right sides of production rules do not have strings of length at least 2 containing terminals.

$$S \to B_0 AS \mid 0$$
$$A \to SB_1 A \mid SSSS \mid B_1 B_0$$
$$B_0 \to 0$$
$$B_1 \to 1$$

And in the second pass assure that the lengths of all right sides of productions do not exceed 2.

$$S \to B_0 X_0 \mid 0$$
$$A \to SX_1 \mid SX_2 \mid B_1 B_0$$
$$B_0 \to 0$$
$$B_1 \to 1$$
$$X_0 \to AS$$
$$X_1 \to B_1 A$$
$$X_2 \to SX_3$$
$$X_3 \to SS$$

2. The CONJECTURE is true. If L is regular and $\varepsilon \notin L$, there is a CFG G in Chomsky Normal Form such that $L = L(G)$. Choose any $w \in L$. If $|w| = 1$, then there must be a production $S \to w$ in G, and thus a derivation of length 1. If $|w| > 1$, we note that if $|\beta| > |\alpha|$ for any step in a derivation $S \overset{*}{\to} \alpha \to \beta \overset{*}{\to} w$, then $|\beta| = |\alpha| + 1$ and this can only happen if we use a production of the form $A \to BC,\ A, B, C \in N$. Starting with S the derivation must use exactly $|w| - 1$ productions of this form. And since all these productions add $|w|$ (not necessarily) distinct nonterminals to the sentential forms of the derivation, we must also replace them with exactly $|w|$ produsctions of the form $A \to a,\ A \in N,\ a \in \Sigma$. Finally this makes exactly $(2 * |w|) - 1$ steps of the derivation.

3. $L = \{0^n 1 0^n 1 0^{2n} \mid n \geq 0\}$ is not context free. Assume that L is context free, and let k be as guaranteed by the Pumping Lemma for context free languages. We know that $0^k 1 0^k 1 0^{2k} \in L$. So the P.L. assures us that it can be expressed as $0^k 1 0^k 1 0^{2k} = uvwxy$ with $|vwx| \leq k$ and $|vx| \geq 1$. Thus, vx is not empty. We now know that vx can not contain 0's from **both** the prefix of k 0's **and** the suffix of $2k$ 0's. So consider uv^2wx^2y. It can't contain 0's from the prefix of k 0's **or** the suffix of $2k$ 0's, or else the suffix wouldn't have

twice as many 0's as the prefix. We also know that vx can't contain any 1's or else uv^2wx^2y would have more than two 1's. So vx must be a nonempty string of 0's, and uv^2wx^2y would increase the number of 0's between the 1's. This is a contradiction, and $\{0^n10^n10^{2n} \mid n \geq 0\}$ can not be context free.