CS5371 Theory of Computation

Homework 2
Due: 2:00 pm, October 31, 2006 (before class)

1. Let G be a CFG in Chomsky normal form that contains b variables. Show that, if G generates some string with a derivation having at least 2^b steps, $L(G)$ is infinite.

2. Give a formal description and the corresponding state diagram of a PDA that recognizes the language $A = \{w \mid 2\#_{\mathtt{a}}(w) \neq 3\#_{\mathtt{b}}(w), w \in \{\mathtt{a},\mathtt{b}\}^*\}$, where $\#_c(w)$ denotes the number of character c occurring in the string w.

3. Let $C = \{xy \mid x, y \in \{\mathtt{0},\mathtt{1}\}^*, |x| = |y|, \text{ and } x \neq y\}$. Show that C is a context-free language.

4. Let $A = \left\{wtw^R \mid w, t \in \{\mathtt{0},\mathtt{1}\}^* \text{ and } |w| = |t|\right\}$. Prove that A is not a context-free language.

5. **(Ogden's Lemma.)** There is a stronger version of the CFL pumping lemma known as *Ogden's lemma*. It differs from the pumping lemma by allowing us to focus on any p "distinguished" positions of a string z and guaranteeing that the strings to be pumped have between 1 and p distinguished positions. The formal statement of Ogden's lemma is: Let L be a context-free language. Then there is a constant p such that for any string z in L with at least p characters, we can mark any p or more positions in z to be distinguished, and then z can be written as $z = uvwxy$, satisfying the following conditions:

   (i) vwx has at most p distinguished positions.

   (ii) vx has at least one distinguished position.

   (iii) For all $i \geq$, uv^iwx^iy is in L.

   Prove Ogden's lemma. (Hint: The proof is really the same as that of the pumping lemma if we pretend that the non-distinguished positions of z are not present as we select a long path in the parse tree for z.)

6. (Bonus Question) In this question, we apply Ogden's lemma and show that the language $L = \left\{\mathtt{a}^i\mathtt{b}^j\mathtt{c}^k \mid i = j \text{ or } j = k \text{ where } i, j, k \geq 0\right\}$ is inherently ambiguous.

   (a) Suppose that $G = (V, T, \Sigma, S)$ is a CFG for L, and let p be the constant specified for G in Ogden's lemma. Assume that $p > 3$.† Consider the string $z = \mathtt{a}^p\mathtt{b}^p\mathtt{c}^{p+p!}$ in L. Suppose we mark all the positions of '$\mathtt{a}$' as distinguished. Let u, v, w, x, y be the five parts of z as specified in the Ogden's lemma. Show that $v = a^t$ and $x = b^t$ for some t.

   (b) Following the above step, show that there exists a variable A in V such that

   $$S \overset{*}{\Rightarrow} uAy \overset{*}{\Rightarrow} uvAxy \overset{*}{\Rightarrow} \mathtt{a}^{p+p!}\mathtt{b}^{p+p!}\mathtt{c}^{p+p!}.$$

   (c) In a similar manner, find another derivation for $\mathtt{a}^{p+p!}\mathtt{b}^{p+p!}\mathtt{c}^{p+p!}$, and show that this derivation corresponds to a distinct parse tree from the derivation in Part (b). Conclude that L is inherently ambiguous.

†We can have this assumption, because Ogden's lemma holds for p implies that it holds for any q with $q > p$.