CSE396 **Problem Set 6** **Due 4/5, 11:59pm**

Reading:

Tuesday's lecture will give examples of using the CFL Pumping Lemma to prove that certain languages are not CFLs. Then it will move on to "Structural Induction." I will really assume that people have read the handout which was mentioned first on Problem Set 5. Then Thursday's lecture will move into Chapter 3 and Turing machines (TMs). Please read both sections 3.1 and 3.2 in one gulp, because there may be time to cover how nondeterministic two-tape Turing machines that obey two certain restrictions give a definition of pushdown automata (PDAs). Hence also skim section 2.2, but ignore the dedicated PDA notation given there.

Assignment 6, due Thu. 5 April under the usual terms. It is also a shorter one. Assignment 7 will be longer.

(1) This problem is "HW6 Online Part" on *TopHat*, worth 20 pts. as before.

(2) (a) In my lectures I ignored the provision in Definition 2.8 in the text that in a Chomsky normal form rule $A \longrightarrow BC$, neither B nor C may be the start variable. Explain briefly how you could "patch" it after-the-fact to comply with the text's extra provision, at the same time "restoring" ϵ if it was in the original language. (6 pts.)

(b) Convert the following CFG into an equivalent one with no nullable variables or unit rules:

$$\begin{array}{ll} S \longrightarrow RC \mid AT & A \longrightarrow \epsilon \mid Aa \\ R \longrightarrow aRb \mid Aa \mid Bb & B \longrightarrow \epsilon \mid Bb \\ T \longrightarrow bTc \mid Bb \mid Cc & C \longrightarrow \epsilon \mid Cc \end{array}$$

Note: This generates the strings in $L(r)$, where $r = a^*b^*c^*$, that do *not* have the numbers of a's, b's, and c's all equal. If you take a regular expression r' for the *complement* of $a^*b^*c^*$, and add a rule $S \longrightarrow S_{r'}$ as if you were carrying out problem (1) here, then you would literally get a context-free grammar G' for the complement of the *non-CFL* language $\{ a^n b^n c^n : n \geq 0 \}$ in Example 2.36 in section 2.3. This exemplifies that unlike the class of regular languages, the class of context-free languages is **not** closed under complements. (12 pts., making 18 on the problem. You need not convert the grammar all the way to Chomsky normal form.)

(3) Let $M = (Q, \Sigma, \delta, s, F)$ be a DFA such that $s \notin F$. Explain how to build a CFG G *in Chomsky normal form* such that $L(G) = L(M)$. You may use the *TopHat* Q8 as a hint, but note that the fact that a DFA "Is-A" GNFA is not enough—you need to build G directly into Chomsky NF and this will need special aspects of M being a DFA. You may suppose $\Sigma = \{0, 1\}$ if you wish, though it is also possible to formulate the construction of G more abstractly for any Σ.

Also say how you would tweak your G into a grammar G' such that the right-hand side of every rule begins with a terminal. That is, G' will be in Greibach normal form as defined on Assignment 5—but you need not look up anything about this form to answer the question. (15 + 3 = 18 pts., for 56 total on the set)