CSE396, Spring 2019 Problem Set 6 Due Fri. 4/5, 11:59pm

Reading. I have decided to skip the full conversion to Chomsky normal form and postpone the step of identifying nullable variables until we hit the algorithms in Chapter 4. The *fact* of the conversion will be used to visualize the proof of the CFL Pumping Lemma using binary trees rather than arbitrary trees as the text uses in section 2.3. For Tuesday, please read that section with attention to how to prove that languages are not CFLs. Sections 2.2 and 2.4 are skipped except that later when we represent deterministic and nondeterministic pushdown automata as special cases of two-tape Turing machines, we will refer back to a few facts (without the proofs) from them. Hence for Thursday, please read ahead to Turing machines in section 3.1.

This assignment is due on **Friday** 4/5, not Thursday. Both it and the coming week's recitations are keyed to the topic of CFG soundness proofs—by structural induction—and a lighter coverage of comprehensiveness. Prelim II is now fixed for **Thursday, April 25**.

Homework—part online and all *individual work*—due **Fri.** 4/5, 11:59pm.

(1) Using *TopHat*, the "Worksheet" titled **Spr'19 HW6.1**. There are 10 questions, each worth 2 points, for 20 total. All are unique-answer questions with 1 attempt given.

(2) Over $\Sigma = \{a, b\}$, define E to be the language of strings that do *not* have bb as a substring. Define the grammar $G = (V, \Sigma, \mathcal{R}, S)$ with variables S, A, B and rules $\mathcal{R} =$

$$\begin{aligned} S &\rightarrow \epsilon \mid b \mid BS \mid SA \\ A &\rightarrow aS \mid AA \\ B &\rightarrow a \mid bAaB \end{aligned}$$

These are the same grammar and language as on TopHat questions 6–10, and it is highly recommended to do those first (including explanations given).

Prove by the given structural induction technique that $L(G) \subseteq E$. (18 pts.)

(3) Let $G = (\{ S, A \}, \Sigma, \mathcal{R}, S)$ be the context-free grammar with $\Sigma = \{ a, b \}$ and rules $\mathcal{R} =$

$$\begin{aligned} S &\rightarrow SS \mid ASa \mid aA, \\ A &\rightarrow bA \mid SA \mid a \end{aligned}$$

(a) Prove by structural induction that every string generated by G (from the start symbol S, that is) has an even number ≥ 2 of a's. Show the analysis of each rule. (18 pts.)

(b) Is the grammar comprehensive for this property? If you say yes, give a parsing algorithm to show it; if you say no, find a string $x \notin L(G)$ that has an even number of a's—and strengthen the "meanings" of the variables used in your proof of (a) to demonstrate clearly that x is excluded by the grammar. (9 pts., = 27 total on this problem and 65 total for the problem set)