CSE396, Spring 2026 **Problem Set 6** **Due Fri. 3/27, 11:59pm**

Reading. The one remaining topic in Section 2.1 is Chomsky normal form. We view it is a lead-in to the **CFL Pumping Lemma** in section 2.3, allowing lectures to do the proof using binary trees rather than arbitrary trees as the text uses in section 2.3. We will *skip* section 2.2 for now, and section 2.4 almost entirely, but when we hit Chapter 3 the week after, we will re-define nondeterministic and deterministic **pushdown automata** as special cases of two-tape Turing machines instead. There is a chance that the Thursday 3/26 lecture will introduce Turing machines in the last 20 minutes or so.

This homework is based on last week's lectures. If you are among the many who skipped the Thursday 3/12 lecture, you should review the video as well as the posted notes, as the video included a demo of C/C++/Java expression syntax that is treated further in TopHat questions 1–5. You can find further more-specific help on that in the notes https://cse.buffalo.edu/˜regan/cse396/CSE396lect033021.pdf from Spring 2019; some of that but not as much will be reviewed in your Tue. 3/24 lecture. TopHat questions 6–10 are on a rearranged version of the grammar that was covered in lectures last week, and this grammar reappears with further questions in the written problem (2) below. Problem (3) is entirely new.

Homework—part online and all *individual work*—due **Fri. 3/27, 11:59pm**.

(1) Using *TopHat*, the "Worksheet" titled **Spr 26 HW6 Online Part**. 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). The problem for written work here is:

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)