Homework 4: Context-free languages

Theory of Computing (CSE 30151), Spring 2023

Due: 2023-02-24 11:59pm

Instructions

- Create a PDF file (or files) containing your solutions. You can write your solutions by hand, but please scan them into a PDF.
- Please name your PDF file(s) as follows to ensure that the graders give you credit for all of your work:
  - If you're making a complete submission, name it `netid-hw4.pdf`, where *netid* is replaced with your NetID.
  - If you're submitting some problems now and want to submit other problems later, name it `netid-hw4-123.pdf`, where `123` is replaced with the problem number(s) you are submitting at this time.
- Submit your PDF file(s) in Canvas.

Problems (10 points each)

1. **Arithmetic expressions.** Consider the grammar G_4 (page 105) for arithmetic expressions, with start symbol E:

$$E \to E \texttt{+} T \mid T$$
$$T \to T \texttt{*} F \mid F$$
$$F \to \texttt{(}E\texttt{)} \mid \texttt{a} \mid \texttt{b} \mid \texttt{c}$$

   (a) [cf. Exercise 2.1] Give derivations for the following strings. Although we didn't cover derivation trees in class, if you prefer to write your derivations as trees, you may.

   i. `a + b + c`
   ii. `a * b + c`
   iii. `a * (b + c)`

(b) Modify G_4 to allow an exponentiation operator $\uparrow$.

- It should have *higher precedence* than multiplication; that is, in the derivation of the string $\texttt{a} * \texttt{b} \uparrow \texttt{c}$, there should be a nonterminal that rewrites to $\texttt{b} \uparrow \texttt{c}$, and there should not be a nonterminal that rewrites to $\texttt{a} * \texttt{b}$.
- It should be (unlike $*$ and $+$) *right-associative*; that is, in the derivation of the string $\texttt{a} \uparrow \texttt{b} \uparrow \texttt{c}$, there should be a nonterminal that rewrites to $\texttt{b} \uparrow \texttt{c}$, and there should not be a nonterminal that rewrites to $\texttt{a} \uparrow \texttt{b}$.

2. Write a PDA for the language

$$L_2 = \{w \in \{\texttt{2}, \overline{\texttt{1}}\}^* \mid w \text{ has twice as many } \overline{\texttt{1}}\text{'s as } \texttt{2}\text{'s}\}.$$

In other words, if we treat $\overline{\texttt{1}}$ as -1, the total of all the symbols should be zero. For example, $\overline{\texttt{1}}\texttt{2}\overline{\texttt{1}} \in L_2$, but $\texttt{2}\overline{\texttt{1}}\texttt{2}\overline{\texttt{1}} \notin L_2$. Please add a brief explanation of why your PDA works.

3. [Exercise 2.6b] Write both a PDA *and* a CFG for the language

$$L_3 = \overline{\{\texttt{0}^n\texttt{1}^n \mid n \geq 0\}}.$$

For example, $\texttt{000111} \notin L_3$. Please add a brief explanation of why they work.

Hint: First prove that this is equal to $\{\texttt{0}^m\texttt{1}^n \mid m \neq n\} \cup \overline{\texttt{0}^*\texttt{1}^*}$.