Homework 3 – Due Friday, September 27, 2019 <u>before</u> noon

Reminder Collaboration is permitted, but you must write the solutions *by yourself without assistance*, and be ready to explain them orally to the course staff if asked. You must also identify your collaborators. Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden.

Exercises Please practice on exercises and solved problems in Chapter 2. The material they cover may appear on exams. Pay especial attention to **Problem 2.18**. It will be helpful for one of the problems on this homework.

Problems

1. (**CFGs**) Give CFGs that generate the following languages. Unless specified otherwise, your CFGs should have at most 2 variables, and the alphabet is $\Sigma = \{0, 1\}$.

   (a) $L_1 = \{w |\ w \in \{\mathtt{0,1}\}^*$ and w contains a number of 1s that is divisible by $3\}$.

   (b) Given a string w over alphabet $\Sigma = \{0, 1\}$ with a zeros and b ones, define $\text{SORT}(w) = \mathtt{0}^a\mathtt{1}^b$. For example, $\text{SORT}(\mathtt{11001}) = \mathtt{00111}$.
   Let $\Sigma_2 = \{\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\}$. For this part, your grammar should have at most 3 variables.
   $L_2 = \big\{w \in \Sigma_2^* \mid$ the bottom row of w is the string obtained by applying the SORT operation to the top row of $w\big\}$.

   (c) $L_3 = \{w |\ w$ is a balanced string of parentheses and brackets$\}$. The alphabet here is $\Sigma = \{\mathtt{(,),[,]}\}$.

   (d) L_4 is the collection of all strings that contain at least one 1 in their second half (if the string is of odd length, we exclude the middle symbol to construct the second "half"). In other words, $L_4 = \{uv \mid u \in \Sigma^*, v \in \Sigma^*\mathtt{1}\Sigma^*$ and $|u| \geq |v|\}$.

2. (**PDAs**) Draw state diagrams of PDAs (with as few states as you can) that recognize the following languages. Write algorithmic descriptions for your PDAs.

   (a) The language L_2 from Problem 1(b): (i) diagram; (ii) algorithmic description.

   (b) The language L_4 from Problem 1(d): (i) diagram; (ii) algorithmic description.

3. (**Non-CFLs**) Prove that the following languages are not context-free. Use the definition of the SORT operation from Problem 1(b).

   (a) The language $A = \{w \circ \text{SORT}(w) \mid w \in \{0, 1\}^*\}$.

   (b) Let $\Sigma_3 = \left\{ \begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix} \right\}$.
   $B = \big\{w \in \Sigma_3^* \mid$ the middle row of w is the reverse of the top row of w **and** the bottom row of w is the string obtained by applying the SORT operation to the top row of $w\big\}$.
   For example, $\begin{bmatrix}1\\1\\0\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\0\\0\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}\begin{bmatrix}1\\0\\1\end{bmatrix}\begin{bmatrix}0\\1\\1\end{bmatrix}$ are in B, but $\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}$ is not.

(c) The following language over the alphabet $\{\mathtt{a}, \mathtt{b}\}$:
$C = \{\mathtt{a}^i\mathtt{b}^j \mid i, j \geq 0 \text{ and if } i = 1 \text{ then } j \text{ is a prime}\}$.
(*Careful:* C satisfies the pumping lemma for CFLs! Make sure you understand why, but you don't need to write it down.)

4* (**Optional, no collaboration is allowed; BALANCE operation**) Given a string w over alphabet $\Sigma = \{0, 1\}$, define

$$\text{BALANCE}(w) = \begin{cases} (01)^x & \text{if } x = y, \\ (01)^y 0^{x-y} & \text{if } x > y, \\ (01)^x 1^{y-x} & \text{if } x < y, \end{cases}$$

where x is the number of 0s in w and y is the number of 1s in w.

For example, BALANCE(000111) = 010101 and BALANCE(0100000111) = 0101010100. Note that w and BALANCE(w) have the same number of 0s and 1s, but in the beginning of the string BALANCE(w), the 0 and 1 characters alternate until one of them runs out.

Define the corresponding operation on languages:

$$\text{BALANCE}(L) = \{\text{BALANCE}(w) \mid w \in L\}.$$

(a) Give BALANCE(L) for the language $L = \{00011, 000, 11110, 101, 110\}$.

(b) Give a regular expression for the language BALANCE(0^*1^*).

(c) Show that the class of regular languages is *not* closed under the BALANCE operation.

(d) Prove that if A is regular then there is a PDA that recognizes BALANCE(A). An algorithmic description of your PDA is sufficient.
Hint: Make sure that your PDA checks two conditions for every string it accepts: (1) that it is of the form specified by your regular expression; (2) the $\mathtt{0}$'s and $\mathtt{1}$'s in the string can be reordered to obtain a string in A.