CSCI 3130: Formal Languages and Automata Theory
The Chinese University of Hong Kong, Fall 2018

Homework 3
due 11:59pm Thursday October 18

Collaborating on homework is encouraged, but you must write your own solutions in your own words and list your collaborators. Copying someone else's solution will be considered plagiarism and may result in failing the whole course.

Please answer clearly and concisely. Explain your answers. Unexplained answers will get lower scores or even no credits.

(1) (30 points) For each of these languages, give both a context-free grammar *and* a pushdown automaton. Briefly explain how your CFG and PDA work; answers without sufficient explanation will get no points. Your CFG should be relatively simple and contain at most 6 variables.

(a) $L_1 = \{x\#y \mid y \text{ is a subsequence of } x^R, \text{ and } x \in \{\mathtt{a},\mathtt{b}\}^*\}$, $\Sigma = \{\mathtt{a},\mathtt{b},\#\}$
Recall that w^R is w written backwards, and u is a subsequence of v if u can be obtained from v by deleting zero or more symbols from v. For example, `ace` is a subsequence of `w`**`a`**`t`**`c`**`h`**`e`**`r`. Formally, $u = u_1u_2\dots u_k$ is a subsequence of v if $v = s_0u_1s_1u_2\dots s_{k-1}u_ks_k$ for some strings $s_0, s_1, \dots, s_k$.

(b) $L_2 = \{\mathtt{0}^i\mathtt{1}^{i+j}\mathtt{2}^j \mid i \geqslant 0, j \geqslant 0\}$, $\Sigma = \{\mathtt{0},\mathtt{1},\mathtt{2}\}$

(c) $L_3 = \{w \in \{\mathtt{a},\mathtt{b}\}^* \mid w \text{ contains at least as many } \mathtt{a}\text{'s as } \mathtt{b}\text{'s}\}$

(2) (20 points) Consider the following context-free grammar G that describes simple mathematical expressions involving the operators $\star$ (multiplication), ^ (exponentiation), and variable `x`:

$$E \to E{\star}E \mid E\verb|^|E \mid \mathtt{x}$$

The alphabet of G consists of $\star$, ^, `x`.

(a) Convert G to Chomsky Normal Form.

(b) Apply the Cocke–Younger–Kasami algorithm to obtain a parse tree for the following string: $\mathtt{x}{\star}\mathtt{x}\verb|^|\mathtt{x}$. Show the table of variables that generate every substring. Also draw the parse tree you get.

(c) Give a CFG G' that describes the same language as G but is not ambiguous. (Your G' needs not be in Chomsky Normal Form.)

(3) (40 points) Consider the following languages. For each of the languages, say whether the language is (1) regular, (2) context-free but not regular, or (3) not context-free. Explain your answer (give a DFA or argue why one exists, give a CFG or PDA, apply appropriate pumping lemma or give pairwise distinguishable strings).

(a) $L_1 = \{x\#x^R\#x \mid x \in \{\mathtt{a},\mathtt{b}\}^*\}, \Sigma = \{\mathtt{a},\mathtt{b},\#\}$
Recall that x^R is x reversed.

(b) $L_2 = \{w \in \{\mathtt{a},\mathtt{b},\mathtt{c}\}^* \mid w \text{ contains the same number of } \mathtt{a}\text{'s}, \mathtt{b}\text{'s}, \mathtt{c}\text{'s}\}$

(c) $L_3 = \{w\#w \mid w \in \{\mathtt{a},\mathtt{b}\}^* \text{ is a palindrome}\}$, $\Sigma = \{\mathtt{a},\mathtt{b},\#\}$
Recall that a string w is a palindrome if it reads the same forwards and backwards.

(d) $L_4 = \{w \in \{\mathtt{a}, \mathtt{b}\}^* \mid \text{The number of } \mathtt{a}\text{'s in } w \text{ is exactly three times the number of } \mathtt{b}\text{'s in } w\}$

(4) (10 points) Context-free grammars are sometimes used to model natural languages. In this problem you will model a fragment of the English language using context-free grammars. Consider the following English sentences:

`The girl met the boy.`

`The girl that the teacher knows met the boy.`

`The girl that the teacher that the staff saw knows met the boy.`

`The girl that the teacher that the staff that the reporter interviews saw knows met the boy.`

This is special type of sentences built from a subject (`the girl`), a relative pronoun (`that`) followed by another sentence, a transitive verb (`met`) and an object (`the boy`).

Give a context-free grammar G that models this special type of sentences. Your grammar should generate the above sentences (among others) in lower case letters and without punctuation (`the girl met the boy`).