Each of the problems is worth 10 points. Please write your solutions clearly and concisely. If you do not explain your answer you will be given no credit. You are encouraged to collaborate on the homework, but you *must* write your own solutions and list your collaborators on your solution sheet. Copying someone else's solution will be considered plagiarism and may result in failing the whole course.

Please turn in the solutions by 11.59pm on Thursday 16 October. The homework should be dropped off in the box labeled CSC 3130 on the 9th floor of SHB. Late homeworks will not be accepted.

In the context-free grammar descriptions below, the start variable is always the one on the left-hand side of the first production.

Problem 1

Give context-free grammars for the following languages. Provide a short description of what the variables and rules in your grammar represent.

(a) $L_1 = \{x\#y : x^R \text{ is a substring of } y,\ x, y \in \{a, b\}\}$, $\Sigma = \{a, b, \#\}$. (Recall that x^R is the reverse of x.)

(b) $L_2 = \{a^i b^{2i+1} : i \geq 0\}$,

(c) $L_3 = \{a^m b^n c^p d^q : m + n = p + q\}$, $\Sigma = \{a, b, c, d\}$.

(d) $L_4 = \{a^i b^j : i > j\}$, $\Sigma = \{a, b\}$.

(e) $L_5 = \{x : x \text{ has the same number of } a\text{s and } b\text{s or same number of } b\text{s and } c\text{s.}\}$, $\Sigma = \{a, b, c\}$.

Problem 2

Consider the following context-free grammar G:

$$
\begin{aligned}
S &\to TC \mid AR \\
T &\to aTb \mid \varepsilon \\
R &\to bRc \mid \varepsilon \\
A &\to Aa \mid \varepsilon \\
C &\to Cc \mid \varepsilon
\end{aligned}
$$

(a) Prove that the language of this grammar is

$$L = \{a^i b^j c^k : i = j \text{ or } j = k\}.$$

(You need to show that all strings in L can be derived from S, and no other strings can be derived from S.)

(b) Give a non-empty string that has two different parse trees in this grammar.

(c) Show that the string *abbcc* has a unique parse tree.

Problem 3

Consider the following context-free grammar G:

$$\begin{aligned} S &\to A \mid B \\ A &\to B \mid C \mid aB \mid b \\ B &\to C \\ C &\to B \mid Aa \end{aligned}$$

(a) Eliminate all unit productions from G.

(b) Convert G to Chomsky Normal Form.

(c) Apply the Cocke-Younger-Kasami algorithm to obtain parse trees for the following strings: *aabaa*, *abaaa*. (Show the run of the algorithm.)

Problem 4

Consider the languages described by the following context-free grammars. For each of the languages, say whether the language is regular or not. If you think it is regular, describe a DFA, NFA, or regular expression for it. If not, prove using the pumping lemma. (For DFAs and NFAs, you do not need to draw a diagram with all the states. A clear description of the automaton suffices.)

(a) $S \to APAPA$
$A \to A0 \mid A1 \mid \varepsilon$
$P \to 011$

(b) $S \to aSbS \mid bSaS \mid \varepsilon$

(c) $S \to SE \mid 0 \mid 1$
$E \to SS$

(d) $S \to aSb \mid aaSb \mid \varepsilon$

(e) $S \to AI.D \mid AI$
$I \to 0 \mid 1D$
$D \to 0D \mid 1D \mid \varepsilon$
$A \to + \mid - \mid \varepsilon$