Each of the problems is worth 10 points. Write your name, student ID, and your TA's name on the solution sheet.

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 20 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.

Problem 1

For each of these languages, give a context-free grammar *and* a pushdown automaton. Give a short explanation of how the CFG/PDA works.

(a) $L_1 = \{wyw^R : \text{the length of } y \text{ is even}\}$, $\Sigma = \{\mathtt{a}, \mathtt{b}\}$.
(Recall that w^R is the reverse of w.)

(b) $L_2 = \{w : w \text{ has the same number of } \mathtt{a}\text{s as } \mathtt{b}\text{s and } \mathtt{c}\text{s together}\}$, $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$.
(The CFG for this part is optional.)

(c) $L_3 = \{\mathtt{a}^i\mathtt{b}^j\mathtt{c}^k : i > j \text{ or } j > k, \text{ where } i, j, k \geq 0\}$, $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$.

(d) (**Extra credit**) $L_4 = \{xy\colon |x| = |y| \text{ and } x \neq y\}$, $\Sigma = \{\mathtt{a}, \mathtt{b}\}$.

Problem 2

Consider the following context-free grammar G that describes (nonempty) regular expressions over the alphabet $\{0, 1\}$:

$$R \to \mathtt{(}R\mathtt{)} \mid R\mathtt{+}R \mid RR \mid R\mathtt{*} \mid \mathtt{0} \mid \mathtt{1} \mid \mathtt{e}$$

The alphabet of G consists of the symbols `(`, `)`, `+`, `*`, `0`, `1`, and `e`. Here `+` and `*` describe the union and star operators, while `e` describes the empty string.

(a) Convert G to Chomsky Normal Form.

(b) Apply the Cocke-Younger-Kasami algorithm (algorithm 2 from lecture 8) to obtain parse trees for the following strings: `(1+0)*`, `0+01`, `(1+e)1*`. Some of these expressions several parse trees; which ones describe the intended meaning of the expression?

(c) Give a CFG G' that describes the same language as G but is not ambiguous. Moreover, each parse tree in G' should describe the intended meaning of the corresponding regular expression.

Problem 3

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. Justify your answer (e.g. give a DFA, a CFG, and/or apply the appropriate pumping lemma, with explanation).

(a) $L_1 = \{\mathtt{a}^n\mathtt{b}^n\mathtt{a}^n\mathtt{b}^n : n \geq 0\}$, $\Sigma = \{\mathtt{a}, \mathtt{b}\}$.

(b) $L_2 = \{w^R\#z : w \text{ is a substring of } z,\ w, z \in \{\mathtt{a}, \mathtt{b}\}^*\}$, $\Sigma = \{\mathtt{a}, \mathtt{b}, \#\}$.

(c) $L_3 = \{w\#z : w \text{ is a substring of } z,\ w, z \in \{\mathtt{a}, \mathtt{b}\}^*\}$, $\Sigma = \{\mathtt{a}, \mathtt{b}, \#\}$.

(d) $L_4 = \{x{+}y{=}z : x + y = z \text{ in unary where } x, y, z \in \mathtt{1}\mathtt{1}^*\}$, $\Sigma = \{\mathtt{1}, +, =\}$.
For example, $\mathtt{1} + \mathtt{11} = \mathtt{111} \in L_4$ but $+\mathtt{1} = \mathtt{1} \notin L_4$, $\mathtt{1} + \mathtt{1} = \mathtt{111} \notin L_4$.

Problem 4

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 is pretty.
The girl that the boy likes is pretty.
The girl that the boy that the cook pushed likes is pretty.
The girl that the boy that the cook that the girl knows pushed likes is pretty.
```

This is a special type of sentence built from a subject (`The girl`), a relative pronoun (`that`) followed by another sentence, a verb (`is`) and an adjective (`pretty`).

(a) Give a context-free grammar G that models this special type of sentence. Your terminals should be words or sequences of words like `pretty` or `the girl`.

(b) Is the language of G regular? If so, write a regular expression for it. If not, prove using the pumping lemma for regular languages.

(c) Can you give an example of a sentence that is in G but does not make sense in common English?