

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 22 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

For each of the following languages, give *both* a context-free grammar and a pushdown automaton. Provide a short description of what the variables and rules in your CFG represent, and a short and clear description of how the PDA works.

- (a) $L_1 = \{ywy^R : y, w \in \Sigma^*\}, \Sigma = \{a, b\}$.
- (b) $L_2 = \{a^{2i}b^{3i} : i \geq 0\}, \Sigma = \{a, b\}$.
- (c) $L_3 = \{x : x \text{ has fewer } a\text{s than } b\text{s and } c\text{s together}\}, \Sigma = \{a, b, c\}$.
- (d) $L_4 = \{a^ib^jc^k : i = j \text{ or } j = k \text{ or } i = k\}, \Sigma = \{a, b, c\}$.

Problem 2

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

$$R \rightarrow (R) \mid R+R \mid RR \mid R^* \mid 0 \mid 1 \mid \mathbf{e}$$

The alphabet of G consists of the symbols $(,), +, *, 0, 1,$ and \mathbf{e} . Here $+$ and $*$ describe the union and star operators, while \mathbf{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: $(0+1)^*, \mathbf{e} + 0 + 1, \mathbf{e} + 01$. 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. Explain your answer (e.g. give a DFA or argue why one exists, give a CFG, apply the appropriate pumping lemma).

- (a) $L_1 = \{xwywz : w \text{ has length at least } 2, x, y, z, w \in \Sigma^*\}, \Sigma = \{a, b\}$.
- (b) $L_2 = \{xyx^Ry^R : x, y \in \Sigma^*\}, \Sigma = \{a, b\}$.
- (c) $L_3 = \{a^ib^{2i}c^{3i} : i \geq 0\}, \Sigma = \{a, b, c\}$.
- (d) $L_4 = \{x\#y\#z : y^R \text{ is a substring of } xz; x, y, z \in \{a, b\}^*\}, \Sigma = \{a, b, \#\}$.
- (e) $L_5 = \{x\#y\#z : y^R \text{ is a substring of both } x \text{ and } z; x, y, z \in \{a, b\}^*\}, \Sigma = \{a, b, \#\}$.

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 clerk pushed likes is pretty.

The girl that the boy that the clerk 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?