Homework 4: Context-free Grammars

CSE 30151 Spring 2022

Due: Thursday, February 24 at 11:59pm

Instructions

- Use this document to create a PDF file containing your solutions. Do this either by (1) printing this document, writing your solutions on it, and scanning[1] your work into a PDF; or (2) writing your solutions on this PDF digitally. Either way, please ensure that your work is clearly legible.
- If you need extra blank pages, feel free to insert them as needed. The amount of blank space provided beneath a problem does not necessarily indicate the expected length of your solution.
- You have the option to submit your solutions all at once or in parts; late penalties will only be applied to problems that are late. Individual problems cannot be submitted for grading in this way more than once.
- If you plan to submit some parts of your assignment late, before the deadline, upload a single PDF containing the problems you have completed so far. Do not include solutions to problems you want graded later for late credit. After the deadline, if you want to submit additional problems, add them to your original PDF and upload it again.[2]
- Submit your PDF file in Canvas under Assignments > Homework 4: Context-free Grammars. You may re-submit your work any number of times before the due date.

[1] For tips on scanning your work using your mobile device, see `https://help.gradescope.com/article/0chl25eed3-student-scan-mobile-device`.

[2] For tips on concatenating your old and new PDFs together, see `https://help.gradescope.com/article/tp9kl4yx4q-student-troubleshooting-submissions`.

1. Write CFGs generating each of the following languages.

   (a) $L_1 = \{w \in \{\mathtt{a}, \mathtt{b}\}^* \mid w = w^R\}$ **(1 point)**

   (b) $L_2 = \{\mathtt{a}^n\mathtt{b}^m \mid 0 \leq n \leq m \text{ and } m - n \text{ is even}\}$ **(1 point)**

   (c) $L_3 = \{\mathtt{a}^i\mathtt{b}^j\mathtt{c}^k \mid i, j, k \geq 0 \wedge (i \geq k \vee j \geq k)\}$ **(2 points)**

(d) $L_4 = \{\mathtt{a}^n\mathtt{b}^m \mid 2n = 3m + 1\}$ **(2 points)**

(e) $L_5 = \{\mathtt{a}^n\mathtt{b}^m \mid 2n \neq 3m + 1\}$ **(2 points)**

2. Consider the following CFG for a fragment of English:

$$\begin{aligned}
\langle \text{SENTENCE} \rangle &\rightarrow \langle \text{NOUN-PHRASE} \rangle \langle \text{VERB-PHRASE} \rangle \\
\langle \text{NOUN-PHRASE} \rangle &\rightarrow \langle \text{DETERMINER} \rangle \langle \text{NOUN} \rangle \\
\langle \text{VERB-PHRASE} \rangle &\rightarrow \langle \text{VERB} \rangle \langle \text{NOUN-PHRASE} \rangle \\
\langle \text{DETERMINER} \rangle &\rightarrow \texttt{a} \mid \texttt{the} \\
\langle \text{NOUN} \rangle &\rightarrow \texttt{boy} \mid \texttt{girl} \mid \texttt{flower} \\
\langle \text{VERB} \rangle &\rightarrow \texttt{likes} \mid \texttt{sees}
\end{aligned}$$

(a) In English, a verb must match its subject in grammatical number, which is either singular or plural. Modify the CFG above so that it supports both singular and plural nouns, making sure that subjects and their verbs always agree in grammatical number. Your grammar should generate the following sentences:

the boy sees the girls
girls like the flower
a boy sees the flowers

Your grammar should *not* generate the following sentences (here, an asterisk * is the standard notation in linguistics for an ungrammatical sentence; it is not part of the string):

*the boy sees a girls
*girls likes the flower
*a boy see the flowers

(4 points)

(b) The original CFG above does not contain any recursion, so it generates a finite language. One recursive structure in English is the attachment of relative clauses to noun phrases (e.g. "the boy who likes the girl who likes the flower"). Modify the CFG above (the original one, not your solution to the previous problem) so that it supports relative clauses. Your grammar should generate the following sentences:

the boy sees the girl who likes the flower
the girl who the boy sees likes the flower
the boy who likes the flower sees a girl who likes the flower
the boy who likes the girl who likes the flower sees a flower
the girl who the boy who likes the flower sees likes the flower

Your grammar should *not* generate the following sentences:

*the boy sees the girl who a boy likes the flower
*the girl who the boy sees the girl likes the flower

(4 points)

(c) **Extra credit** (cannot be submitted late). Combine the modifications from your previous answers into a single CFG that supports both plural nouns and relative clauses, ensuring that the subject and verb of a relative clause always agree in number. **(4 points)**

3. Consider the following fragment of a CFG for a programming language:

$$
\begin{aligned}
\langle \text{PROGRAM} \rangle &\to \langle \text{STATEMENT} \rangle \\
\langle \text{STATEMENT} \rangle &\to \langle \text{EXPRESSION} \rangle \\
&\quad \mid \texttt{if (} \langle \text{EXPRESSION} \rangle \texttt{) then} \langle \text{STATEMENT} \rangle \\
&\quad \mid \texttt{if (} \langle \text{EXPRESSION} \rangle \texttt{) then} \langle \text{STATEMENT} \rangle \texttt{ else} \langle \text{STATEMENT} \rangle \\
&\quad \mid \texttt{for (} \langle \text{EXPRESSION} \rangle \texttt{ ;} \langle \text{EXPRESSION} \rangle \texttt{ ;} \langle \text{EXPRESSION} \rangle \texttt{)} \langle \text{STATEMENT} \rangle \\
\langle \text{EXPRESSION} \rangle &\to \texttt{id} \mid \cdots \\
&\quad \vdots
\end{aligned}
$$

Show that this grammar is ambiguous, and modify the grammar (without modifying the language it generates) to remove the ambiguity. **(4 points)**

4. Convert the following CFG to Chomsky normal form. **(4 points)**

$S \to \mathtt{a}S\mathtt{a} \mid B$
$B \to \mathtt{bbb}C$
$C \to \mathtt{c}C \mid \varepsilon$

Changelog

- **Feb 23**: Updated submission instructions.