# HW2 - Formal grammars (Solution)

March 17, 2025

1. What are the languages engendered by the following context-free grammars:
   (a) *[1.5 pts.]* $\mathcal{G}_1 = (\{a, b\}, \{S\}, S, \{S \to aSa \mid bSb \mid \varepsilon\})$

   ***Solution:*** $L(\mathcal{G}_1) = \{ww^{\mathcal{R}} \mid w \in \{a, b\}^*\}$.

   (b) *[2.5 pts.]* $\mathcal{G}_2 = (\{a, b\}, \{S, T\}, S, \{S \to aSa \mid bSb \mid aTb \mid bTa, T \to Ta \mid Tb \mid \varepsilon\})$

   ***Solution:*** $L(\mathcal{G}_2) = \{w \in \{a, b\}^* \mid w \neq w^{\mathcal{R}}\}$.

2. Give a context-free grammar for each of the following languages:
   (a) *[2.5 pts.]* $\{a^m b^n c^p \mid n \geq m + p\}$

   ***Solution:***
   $$\begin{array}{lcl} S & \to & ABC \\ A & \to & aAb \mid \varepsilon \\ B & \to & bB \mid \varepsilon \\ C & \to & bCc \mid \varepsilon \end{array}$$

   (b) *[1 pts.]* $\{va^n \mid v \in \{a, b\}^n \;\wedge\; |v|_a = n\}$

   ***Solution:*** $S \to aSa \mid \varepsilon$

   **There was an error in the question; the intended language was**

   $\{va^n \mid v \in \{a, b\}^* \;\wedge\; |v|_a = n\}$
   $$\begin{array}{lcl} S & \to & BaSa \mid B \\ B & \to & bB \mid \varepsilon \end{array}$$

   **For those who tried (and almost got it right), this question awarded up to 2.5 points.**

   (c) *[2.5 pts.]* $\{w \in \{a, b\}^* \mid |w|_a \geq |w|_b\}$

   ***Solution:*** $S \to aSbS \mid bSaS \mid aS \mid \varepsilon$

3. *[5 pts.]* Propose a context-free grammar that generates the set of valid arithmetic expressions (in $\mathbb{N}$) over $\{0, 1, \ldots, 9, \times, +, (, )\}$ with the two following restrictions:
   - the numbers should not contain any leading 0's, except for 0 itself, and
   - there should not be any redundant parentheses.

   For instance, expressions in (1) are valid, while those in (2) are not.

   (1) a. $1234 \times 0 \times (1 + 654 + 292929)$
       b. $4 \times 911 + 2 \times (42 + 5555 \times 886) \times 1$
       c. $9 \times (1 + 2 + 7) \times (4 \times 2 + 1) + 5 \times 112$
       d. $(0 + 4) \times (5 \times 7 + 91) + 4$

   (2) a. $*\,010$
       b. $*\,1 \times (2)$
       c. $*\,4 \times 911 + (2 \times 42 + 5555)$
       d. $*\,9 \times (1 \times 2) + 74$

   Is your grammar ambiguous? If so, show two parse trees for a valid expression; otherwise, explain informally why your grammar is not (a formal proof is not expected).

***Solution:***
$$\begin{array}{lcl} E & \to & E + M \mid M \\ M & \to & M \times N \mid (A) \times M \mid M \times (A) \mid (A) \times (A) \mid N \\ A & \to & E + M \\ N & \to & 0 \mid 1N' \mid 2N' \mid \cdots \mid 9N' \\ N' & \to & 0N' \mid 1N' \mid 2N' \mid \cdots \mid 9N' \mid \varepsilon \end{array}$$

This grammar is ambiguous due to the production rules for $M$. For instance, the expression $(1+2) \times (2+0) \times 1$ has two derivations according to when '$\times 1$' is derived.

4. *[5 pts.]* Define a non-trivial context-free grammar that could generate the sentences in (3).

(3) a. Alice eats a sweet.
    b. The waiter gives Alice cakes in a tray.
    c. The tabby cat with a grin disappears.
    d. Alice likes that her brother plays football with his friends.

***Solution:***

| | | |
|---|---|---|
| S | $\rightarrow$ | NP VP |
| NP | $\rightarrow$ | PropN |
| | $\mid$ | Det N' |
| | $\mid$ | N' |
| N' | $\rightarrow$ | Adj N' |
| | $\mid$ | N' PP |
| | $\mid$ | N |
| PP | $\rightarrow$ | P NP |
| VP | $\rightarrow$ | VP PP |
| | $\mid$ | $\mathrm{V_t}$ NP |
| | $\mid$ | $\mathrm{V_d}$ NP NP |
| | $\mid$ | $\mathrm{V_t}$ *that* S |
| | $\mid$ | $\mathrm{V_i}$ |
| PropN | $\rightarrow$ | *Alice* |
| Det | $\rightarrow$ | *a* $\mid$ *the* $\mid$ *her* $\mid$ *his* |
| N | $\rightarrow$ | *sweet* $\mid$ *waiter* $\mid$ *cakes* $\mid$ *tray* $\mid$ *cat* $\mid$ *grin* $\mid$ *brother* $\mid$ *football* $\mid$ *friends* |
| Adj | $\rightarrow$ | *tabby* |
| P | $\rightarrow$ | *in* $\mid$ *with* |
| $\mathrm{V_t}$ | $\rightarrow$ | *eats* $\mid$ *likes* $\mid$ *plays* |
| $\mathrm{V_d}$ | $\rightarrow$ | *gives* |
| $\mathrm{V_i}$ | $\rightarrow$ | *disappears* |

Note that the rules N' $\rightarrow$ Adj N' | N' PP allow for the ambiguity in (3c):

- Either there is only one *tabby cat* with a grin: N' $\rightarrow$ Adj N' $\rightarrow$ Adj N' PP, or
- there is only one *cat with a grin* that is tabby: N' $\rightarrow$ N' PP $\rightarrow$ Adj N' PP