NTIN071 A&G: Tutorial 6 – Formal grammars, regular and context-free grammars

Teaching goals: The student is able to

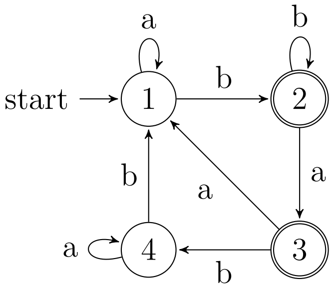
- explain the formal definition of a grammar, and the language it generates,
- give definitions and examples grammars of all types in the Chomsky hierarchy,
- describe a language generated by a given context-free grammar,
- construct a grammar for a language given in set notation,
- convert a finite automaton to a right-linear grammar,
- convert a right-linear grammar to a finite automaton,
- design algorithms to test basic properties of context-free grammars.

In-class problems

Problem 1 (Constructing grammars)**.** Design grammars (of the highest possible type) which generate the following languages ($\Sigma = \{a, b\}$ unless specified otherwise):

(a) $L = \{w \in \Sigma^* \mid |w|_b \text{ is even}\}$
(b) $L = \{ww^R \mid w \in \Sigma^*\}$
(c) $L = \{a^i b^j c^k \mid i = j \text{ or } j = k\}$
(d) $L = \{a^i b^j \mid 0 \leq i \leq j \leq 2i\}$
(e) $L = \{uabbav \mid u, v \in \Sigma^* \text{ and } |u| = |v|\}$
(f) $L = \{w \in \Sigma^* \mid -1 \leq |w|_a - |w|_b \leq 1\}$

Problem 2 (FA to grammar)**.** For the following automaton, find an equivalent grammar. Which class of the Chomsky hierarchy does it belong to?



Problem 3 (Regular grammar to FA)**.** Convert the following right-linear grammar to a finite automaton: $G = (\{S, A, B, C\}, \{a, b\}, \mathcal{P}, S)$ where $\mathcal{P}$ consists of the following:

$$
\begin{aligned}
S &\to abS \mid babA \mid \epsilon \\
A &\to abA \mid aB \mid bC \\
B &\to abS \mid B \mid bC \mid \epsilon \\
C &\to aab \mid A \mid aA \mid \epsilon
\end{aligned}
$$

Problem 4 (Testing properties of context-free languages)**.** Design an (efficient) algorithm which decides whether a given context-free grammar satisfies the given property:

(a) $L(G) \neq \emptyset$
(b) $\epsilon \in L(G)$
(c) $L(G)$ is a finite language

Extra Practice and Thinking

Problem 5 (Constructing grammars)**.** Design grammars (of the highest possible type) which generate the following languages ($\Sigma = \{a, b\}$ unless specified otherwise):

(a) $L = \Sigma^*$
(b) $L = \{a^{2i}b^j \mid i \leq j\}$
(c) $L = \{w \in \Sigma^* \mid |w|_a = 2|w|_b\}$
(d) $L = \{uabbav \mid u, v \in \Sigma^* \text{ and } |u| \neq |v|\}$
(e) $L = \{w\#s^R \mid w, s \in \Sigma^* \text{ and } s \text{ is a subword of } w\}$

Problem 6 (Small grammars generating large (finite) languages)**.** Find a sequence of context-free grammars $G_1, G_2, G_3, \ldots$ (over a given alphabet Σ) such that G_n generates exactly all words of length $\leq 2^n$ (and no other words), and the size of G_n (for simplicity, say the number of symbols in bodies of production rules) is in $O(n)$.

Problem 7 (Context-sensitive Grammar?)**.** Let $G = (\{S, A, B, C\}, \{a, b, c\}, \mathcal{P}, S)$, where:

$$\mathcal{P} = \{S \to aSBC \mid aBC, B \to BBC, C \to CC, CB \to BC,$$
$$aB \to ab, bB \to bb, bC \to bc, cC \to cc\}$$

What language does it generate? Is the grammar G context-sensitive? If not, find an equivalent context-sensitive grammar.