What is a Formal Grammar?

Arno Pauly

January 29, 2021

Defining formal grammars

- We have our alphabet Σ, the elements of which we call *terminal* symbols (and often use $\{a, b, c, \ldots\}$).
- We have a disjoint set of symbols $\mathcal{N}$, the elements of which we call *non-terminal* symbols (and often use capital letters).
- There is a special start symbol $S \in \Gamma$.
- A grammar then is a finite list of pairs (u, w) with $u, w \in (\Sigma \cup \mathcal{N})^*$ (called *production rules*, and often written $u \to w$).

How a formal grammar defines a language

Definition

A grammar G defines a language $L(G) \subseteq \Sigma^*$ by saying that $t \in L(G)$ if we can reach t with the following process:

1. Start with S.
2. Write our current word as $v_0 u v_1$, and pick a rule (u, w). Then replace the current word by $v_0 w v_1$.
3. If the current word is t, stop, else repeat Step 2.

Example

Example

Let $\Sigma = \{a, b\}$, $\mathcal{N} = \{S\}$ and the rules be $S \to \varepsilon$, $S \to aSa$ and $S \to bSb$.

- This grammar defines the palindroms over $\{a, b\}$.

A “real” grammar

Example

Let $\Sigma = \{\text{the}, \text{dog}, \text{cat}, \text{eats}, \text{sleeps}\}$,

$$\mathcal{N} = \{S, \text{NOUN}, \text{NP}, \text{VP}, \text{TRANS-VERB}, \text{INTRANS-VERB}\}$$

and the rules be $S \to$ NP VP, NP $\to$ the NOUN, NOUN $\to$ cat, NOUN $\to$ dog, VP $\to$ INTRANS-VERB, VP $\to$ TRANS-VERB NP, TRANS-VERB $\to$ eats, INTRANS-VERB $\to$ sleeps, INTRANS-VERB $\to$ eats.

Task

Find all “words” (ie sentences) belonging to the language of this grammar.

Outlook

- Without restrictions on how rules might look like, it can be very time consuming to show that a word belongs to the language,
- and impossible(!!) to show that a word does not.
- We can formalize the derivation process a bit more, by introducing the *transitive closure* of a relation.