

Grammatica

Una grammatica è una quadrupla $G = \langle N, \Sigma, P, S \rangle$, dove

- $N \neq \emptyset$ è un insieme finito di *simboli non terminali*;
- Σ è un alfabeto di *simboli terminali*;
- $P \subseteq (N \cup \Sigma)^+ \times (N \cup \Sigma)^*$ è un insieme finito di *produzioni*;
- $S \in N$ è il *simbolo distinto* (o *iniziale*).

Convenzioni:

$$\begin{array}{ll} A, B, \dots \in N & (\alpha, \beta) \in P \\ a, b, \dots \in \Sigma & X, Y, \dots \in N \cup \Sigma \\ \alpha, \beta, \dots \in (N \cup \Sigma)^* & x, y, \dots \in \Sigma^* \end{array}$$

Le grammatiche si classificano sulla base della gerarchia di Chomsky. Di particolare interesse per noi sono le grammatiche libere da contesto (forma delle produzioni (A, β)) e quelle regolari o lineari destre (produzioni del tipo (A, b) e (A, bA)).

Le grammatiche possono essere espresse in una notazione alternativa detta BNF (Backus-Naur form).

$$\begin{array}{l} E ::= E + T \mid T \\ T ::= T * F \mid F \\ F ::= a \mid \dots \mid z \mid (E) \end{array}$$

è equivalente a

$$\begin{array}{l} \langle \{E, T, F\}, \{a, \dots, z, +, *, (,)\}, \\ \{(E, E + T), (E, T), (T, T * F), (T, F), (F, a), \dots, (F, z), (F, (E))\}, E \rangle \end{array}$$

Il linguaggio generato dalla grammatica G è $L(G) = \{w \in \Sigma^* \mid S \rightarrow^* w\}$. Due grammatiche sono equivalenti se $L(G_1) = L(G_2)$.