Grammatiken

beschreiben Sprachen $L \subseteq \Sigma^*$, indem sie eine Methode beschreiben, mit der genau alle $x \in L$ erzeugt werden können.

z. B.: korrekte HTML, Java, C^{++}, … Programme werden (fast vollständig) durch sog. kontextfreie Grammatik beschrieben.

Grammatiken, Definitionen

Eine **Grammatik** G (vom Typ **Chomsky-0, Rewritingsystem**, **Semi-Thue-System**) ist beschrieben durch ein 4-Tupel (V, Σ, P, S).

- V ist endliches Alphabet von **Variablen**
- Σ ist endliches Alphabet von **Terminalen**
- $S \in V$ ist das **Startsymbol**
- $P \subseteq ((V \cup \Sigma)^+ - \Sigma^*) \times (V \cup \Sigma)^*$ ist eine endliche Menge von **Produktionen** oder **(Ersetzungs-) Regeln**

Für eine Regel $(u,v) \in P$ schreiben wir $u \rightarrow v$.

Beispiel:

$G_2 = (V, \Sigma, P, S)$ mit $V = \{S, B, C\}$,
$\Sigma = \{a, b, c\}$,

$P = \{1: S \rightarrow aSBC,\ 2: S \rightarrow aBC,\ 3: CB \rightarrow BC,$
$4: aB \rightarrow ab,\ 5: bB \rightarrow bb,\ 6: bC \rightarrow bc,$
$7: cC \rightarrow cc\}$

Ableitungen in G

HEINZ NIXDORF INSTITUT
Universität Paderborn
Algorithmen und Komplexität

Def.: $w' \in (V \cup \Sigma)^*$ ist aus $w \in (V \cup \Sigma)^+$ **direkt ableitbar**, ($w \rightarrow w'$), falls Regel $u \rightarrow v \in P$ und $\alpha, \beta \in (V \cup \Sigma)^*$ existieren mit:

$w = \alpha\, u\, \beta,\ w' = \alpha\, v\, \beta$

- w' ist aus w **ableitbar**, falls es $w_0, \ldots, w_l$ gibt mit $w = w_0 \rightarrow w_1 \rightarrow w_2 \rightarrow \ldots\ w_l = w'$ ($w \xrightarrow{*} w'$)

- $L(G) := \{w \in \Sigma^*,\ S \xrightarrow{*} w\}$

Beispiel G_2:

$L(G_2) = \{a^n b^n c^n, n \geq 1\}$

Beispiel:

$$G_3 = (V, \Sigma, P, \langle Ausdruck \rangle), \text{ wobei :}$$

$$V = \{\langle Ausdruck \rangle, \langle Term \rangle, \langle Faktor \rangle\}$$

$$\Sigma = \{a, (,), *, +\}$$

$$P = \begin{cases} 1. & \langle Ausdruck \rangle & \rightarrow & \langle Term \rangle \\ 2. & \langle Ausdruck \rangle & \rightarrow & \langle Ausdruck \rangle + \langle Term \rangle \\ 3. & \langle Term \rangle & \rightarrow & \langle Faktor \rangle \\ 4. & \langle Term \rangle & \rightarrow & \langle Term \rangle * \langle Faktor \rangle \\ 5. & \langle Faktor \rangle & \rightarrow & a \\ 6. & \langle Faktor \rangle & \rightarrow & (\langle Ausdruck \rangle) \end{cases}$$

$L(G_3) =$ "korrekte Klammerausdrücke mit Variable a und Operationen $*, +$"

Die Chomsky-Hierarchie

- Eine Grammatik heißt **kontextsensitiv** oder vom Typ **Chomsky-1**, falls für jede Regel $u \rightarrow v$ gilt: $|u| \leq |v|$.
- Sie heißt **kontextfrei** oder vom Typ **Chomsky-2**, falls alle Regeln vom Typ $u \rightarrow v$ mit $u \in V$ sind.
- Sie heißt **regulär** oder vom Typ **Chomsky-3**, falls alle Regeln vom Typ $u \rightarrow v$ mit $u \in V$ und

  $v \in \{\varepsilon\} \cup \Sigma \cup (\Sigma \times V)$ sind.

(Eine allgemeine Grammatik ist vom Typ **Chomsky-0** .)

Satz: L ist genau dann rekursiv aufzählbar, wenn
es Chomsky-0-Grammatik G gibt mit L = L (G).

Satz: Kontext-sensitive Sprachen sind entscheidbar.