Informatik IV

— Theoretische Informatik —

Inhalt:

1. Formale Sprachen und Automaten
2. Berechenbarkeit
3. Komplexitätstheorie
4. Effiziente Algorithmen und Datenstrukturen

Literatur:

- M. Broy: Informatik, Band 2, Springer Verlag
- U. Schöning: Theoretische Informatik – kurzgefaßt, Spektrum Verlag
- I. Wegener: Theoretische Informatik, Teubner
- Aho, Hopcroft, Ullman: Introduction to automata theory, languages, and computation, Addison-Wesley, 1979

Informationen zu Vorlesung und Übungsbetrieb:

http://www14.in.tum.de/lehre/2000SS/info4/

Ein erstes Beispiel:

<Satz> $\rightarrow$ <Subjekt> <Prädikat> <Objekt>

<Subjekt> $\rightarrow$ <Artikel> <Attribut> <Substantiv>

<Artikel> $\rightarrow$ ε

<Artikel> $\rightarrow$ <der>

<Artikel> $\rightarrow$ <die>

<Artikel> $\rightarrow$ <das>

<Attribut> $\rightarrow$ ε

<Attribut> $\rightarrow$ <Adjektiv>

<Attribut> $\rightarrow$ <Adjektiv> <Attribut>

<Adjektiv> $\rightarrow$ klein

<Adjektiv> $\rightarrow$ groß

u.s.w.

Definition. Eine *Grammatik* ist ein 4-Tupel $G = (V, \Sigma, P, S)$ das folgende Bedingungen erfüllt:

- V ist eine endliche Menge, die Menge der *Variablen*.
- Σ ist eine endliche Menge, das *Terminalalphabet*, wobei $V \cap \Sigma = \emptyset$.
- P ist die Menge der *Produktionen* oder *Regeln*. P ist eine endliche Teilmenge von $(V \cup \Sigma)^+ \times (V \cup \Sigma)^*$. (Schreibweise: $(u, v) \in P$ schreibt man meist als $u \to v$.)
- $S \in V$ ist die *Startvariable*.

Seien $u, v \in (V \cup \Sigma)^*$. Wir definieren die Relation $u \Rightarrow_G v$ (in Worten: u geht unter G unmittelbar über in v), falls u und v die Form haben

$$\begin{aligned} u &= xyz \\ v &= xy'z \quad \text{mit } x, z \in (V \cup \Sigma)^* \text{ und} \\ y &\to y' \quad \text{eine Regel in } P \text{ ist.} \end{aligned}$$

(Bem: Falls klar ist, welche Grammatik gemeint ist, so schreiben wir oft auch einfach kurz $u \Rightarrow v$ anstelle von $u \Rightarrow_G v$.)

Die von G definierte (erzeugte, dargestellte) *Sprache* ist

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

wobei $\Rightarrow_G^*$ die reflexive und transitive Hülle von $\Rightarrow_G$ ist.

Eine Folge von Worten $(w_0, w_1, \ldots, w_n)$ mit $w_0 = S$, $w_n \in \Sigma^*$ und $w_i \Rightarrow w_{i+1}$ für $i = 0, \ldots, n-1$ heißt *Ableitung* von w_n.

Definition. Jede Grammatik ist automatisch vom Typ 0. (D.h., bei Grammatiken von Typ 0 gibt es keinerlei Einschränkungen an die Regeln in P.)

Eine Grammatik ist vom *Typ 1* oder *kontextsensitiv*, falls für alle Regeln $u \to v$ in P gilt: $|u| \leq |v|$.

Eine Grammatik ist vom *Typ 2* oder *kontextfrei*, falls für alle Regeln $u \to v$ in P gilt: $u \in V$ (d.h., u ist eine einzelne Variable) und $|v| \geq 1$.

Eine Grammatik ist vom *Typ 3* oder *regulär*, falls für alle Regeln $u \to v$ in P gilt: $u \in V$ (d.h., u ist eine einzelne Variable) und $w \in \Sigma \cup \Sigma V$ (d.h., v ist entweder eine einzelnes Terminalzeichen oder ein Terminalzeichen gefolgt von einer Variablen).

Eine Sprache $L \subseteq \Sigma^*$ heißt vom Typ i, $i \in \{0, 1, 2, 3\}$, falls es eine Grammatik vom Typ i gibt mit $L(G) = L$.

ε-Sonderregelung: Wegen $|u| \leq |v|$ kann das leere Wort bei Typ 1,2,3 Grammatiken nicht erzeugt werden. Wir erlauben daher die folgende Sonderregelung: Ist $\varepsilon \in L(G)$ erwünscht, so ist die Regel $S \to \varepsilon$ zugelassen, falls die Startvariable S auf keiner rechten Seite einer Produktion vorkommt.