5.4 Grammatiken

Wählen Sie aus jeder Spalte ein Teil: *Mein Hund* | *Mein kleiner Bruder* | *Der Drucker* — *hat* — *meine Hausaufgaben* | *mein Referat* — *gefressen* | *gelöscht* | *versteckt*. Das ergibt $3 \cdot 2 \cdot 3 = 18$ Ausreden, ohne dass Sie eine einzige auswendig lernen. So arbeitet eine Grammatik: wenige Regeln, viele Sätze. Ob die Ausrede glaubwürdig ist, entscheidet sie nicht — das ist Semantik.

Nicht jede Folge von Wörtern ergibt einen Satz:

- Hund alle sind.
- Der Hund ist ein Insekt.
- `while do end x < 3 begin`

Die erste und die dritte Zeile verstoßen gegen die Form, die zweite nur gegen den Inhalt. Für die Form zuständig ist die **Grammatik**. Sie hat zwei Aufgaben: die zulässigen Zeichenfolgen zu beschreiben und die Regeln dafür darzustellen. Die Semantik spielt dabei keine Rolle.

Definition: Grammatik

Eine **Grammatik** ist ein **4-Tupel** $G = (V_N, V_T, P, S)$ mit

- der Menge V_N der **Nichtterminalzeichen** (syntaktische Variablen),
- der Menge V_T der **Terminalzeichen** (auch Token — die Zeichen, die im fertigen Wort stehen dürfen), wobei $V_N \cap V_T = \emptyset$,
- der Menge P der **Produktionen** (Regeln der Form „links $\rightarrow$ rechts") und
- dem **Startsymbol** $S \in V_N$.

Die Angabe als Tupel ist die übliche Schreibweise und wird im Abitur erwartet. Sie zwingt dazu, nichts zu vergessen: vier Klammern, vier Bestandteile.

Eine **Ableitung** ist eine Folge von Ersetzungen, beginnend beim Startsymbol. Die Menge aller ableitbaren Wörter über V_T heißt die von G **erzeugte Sprache** $L(G)$.

5.4.1 Ein erstes Beispiel

$$V_T = \{\text{Katzen}, \text{Hunde}, \text{schnurren}, \text{fressen}\}, \quad V_N = \{S, \text{Subjekt}, \text{Prädikat}\}$$

$$\begin{aligned}
(P_1) &\quad S \rightarrow \text{Subjekt Prädikat} \\
(P_2) &\quad \text{Subjekt} \rightarrow \text{Katzen} \mid \text{Hunde} \\
(P_3) &\quad \text{Prädikat} \rightarrow \text{schnurren} \mid \text{fressen}
\end{aligned}$$

Der Pfeil bedeutet: Das Nichtterminal links darf durch die rechte Seite ersetzt werden, der Strich | trennt Alternativen. Ableitung von „Hunde schnurren":

$$S \Rightarrow \text{Subjekt Prädikat} \Rightarrow \text{Hunde Prädikat} \Rightarrow \text{Hunde schnurren}$$

Der Satz gehört also zu $L(G)$ — dass Hunde nicht schnurren, ist eine Frage der Semantik und geht die Grammatik nichts an. Grafisch notiert man die Ableitung als **Ableitungsbaum**:

```
          S
         / \
   Subjekt  Prädikat
      |        |
    Hunde   schnurren
```

5.4.2 Zweites Beispiel: Klammerterme

Gesucht ist eine Grammatik für korrekt geschachtelte Klammern wie `(())` oder `((()())())`.

$$V_T = \{(,\ )\}, \qquad V_N = \{S\}, \qquad P = \{\, S \to (S) \mid SS \mid \varepsilon \,\}$$

Ableitung von `(())()`:

$$S \Rightarrow SS \Rightarrow (S)S \Rightarrow ((S))S \Rightarrow (())S \Rightarrow (())(S) \Rightarrow (())()$$

Beachten Sie die Regel $S \to (S)$: Das Nichtterminal steht *zwischen* Terminalzeichen. Diese Bauart wird in Abschnitt 5.10 entscheidend.

5.4.3 Drittes Beispiel: ein Satzmuster

$$V_T = \{\text{der, den, Hund, Briefträger, beißt, schlägt}\}$$
$$V_N = \{S, \text{Subjekt, Prädikat, Objekt, Nomen}\}$$
$$P = \{\, S \to \text{Subjekt Prädikat Objekt},\ \text{Subjekt} \to \text{der Nomen},\ \text{Objekt} \to \text{den Nomen},$$
$$\text{Prädikat} \to \text{beißt} \mid \text{schlägt},\ \text{Nomen} \to \text{Hund} \mid \text{Briefträger} \,\}$$

Arbeitsauftrag 4

12. Leiten Sie mit der letzten Grammatik die Sätze „der Briefträger beißt den Hund" und „der Briefträger schlägt den Briefträger" ab und zeichnen Sie je einen Ableitungsbaum.
13. Wie viele verschiedene Sätze erzeugt diese Grammatik insgesamt? Begründen Sie mit einer Rechnung.
14. Leiten Sie `((()))` und `()()()` mit der Klammergrammatik ab.
15. Geben Sie eine Grammatik an, die alle Wörter aus a und b erzeugt, die mit a beginnen.
16. Erweitern Sie die Satzgrammatik um ein Adjektiv vor dem Nomen.