Formale Sprachen

Chomsky-Hierarchie[1]

Im Jahr 1957 veröffentlichte der amerikanische Sprachwissenschaftler Noam *Chomsky* ein Regelwerk, mit dessen Hilfe sich formale Grammatiken in *vier Klassen* einteilen lassen:[2]

Chomsky
vier Klassen

Typ 0 Phrasenstrukturgrammatik

Jede Grammatik ist per Definition immer auch eine Typ-0-Grammatik. Insbesondere unterliegt die Struktur der Produktionen *keinen weiteren vereinbarten Einschränkungen*.

keinen weiteren vereinbarten Einschränkungen

Typ 1 kontextsensitive Grammatik

Eine Grammatik heißt kontextsensitiv, wenn jede Produktionsregel $l \to r$ entweder die Beziehung $|r| \geq |l|$ erfüllt oder die Form $S \to \varepsilon$ aufweist. Ist die Regel $S \to \varepsilon$ enthalten, so darf S in keiner anderen rechten Seite einer Regel vorkommen.

$|r| \geq |l|$
$S \to \varepsilon$

Typ 2 kontextfreie Grammatik

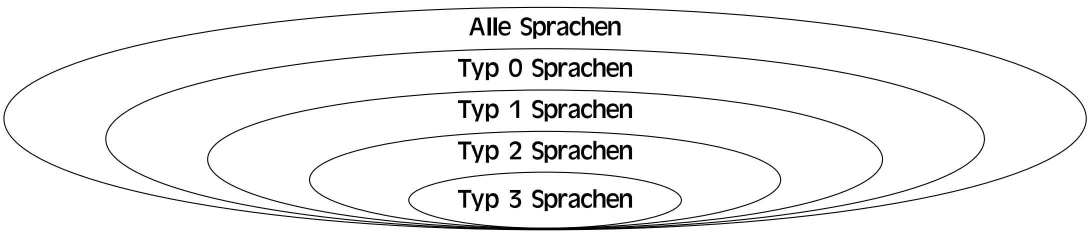
Typ-2-Grammatiken sind dadurch charakterisiert, dass die *linke Seite* einer Produktionsregel ausschließlich aus einer *einzigen Variablen* besteht. Für alle Produktionen $l \to r$ gilt also $l \in V$.

linke Seite
einzigen Variablen

Typ 3 reguläre Grammatik[3]

Reguläre Grammatiken sind kontextfrei und besitzen die zusätzliche Eigenschaft, dass die rechte Seite einer Produktion entweder aus dem *leeren Wort ε* oder einem *Terminalsymbol*, *gefolgt von einem Nonterminal*, besteht. Formal gesprochen besitzt jede Produktion die Form $l \to r$ mit $l \in V$ und $r \in \{\varepsilon\} \cup \Sigma V$.

leeren Wort ε
Terminalsymbol
gefolgt von einem Nonterminal



Eine Sprache L bezeichnen wir als Typ-n-Sprache, wenn eine Typ-n-Grammatik G existiert, die L erzeugt. Die Menge aller Typ-n-Sprachen notieren wir mit dem Symbol $\mathcal{L}_n$. Zwischen den verschiedenen Sprachklassen besteht die folgende Inklusionsbeziehung:

$$\mathcal{L}_0 \supset \mathcal{L}_1 \supset \mathcal{L}_2 \supset \mathcal{L}_3$$

[1] *Wikipedia-Artikel „Chomsky-Hierarchie“*.
[2] Hoffmann, *Theoretische Informatik*, Seite 168.
[3] *Theoretische Informatik – Reguläre Sprachen*, Seite 14.

Beispiel-Sprache je Sprach-Typ

- $L_3 = \{ (ab)^n \mid n \in \mathbb{N} \}$ ist regulär (Typ 3 Sprache)
- $L_2 = \{ a^n b^n \mid n \in \mathbb{N} \}$ ist kontextfrei (Typ 2 Sprache), aber nicht regulär
- $L_1 = \{ a^n b^n c^n \mid n \in \mathbb{N} \}$ ist kontextsensitiv (Typ 1 Sprache), aber nicht kontextfrei
- $L_0 = \{ a^{(2^n)} \mid n \in \mathbb{N} \}$ ist eine Typ 0 Sprache[4]

Terminalsymbol

nicht weiter durch eine Produktionsregel ersetzt werden kann

aus Wörtern durch eine Grammatik neue Wörter bzw. Symbolfolgen produziert

Ein *Terminalsymbol* (auch Terminalzeichen oder kurz Terminal genannt) einer formalen Grammatik ist ein Symbol, das einzeln *nicht weiter durch eine Produktionsregel ersetzt werden kann*.[5]

Eine Produktionsregel (auch Regel, Produktion oder Ersetzungsregel genannt) ist in der Theorie formaler Grammatiken eine Regel, die angibt, wie *aus Wörtern durch eine Grammatik neue Wörter bzw. Symbolfolgen produziert* werden.[6]

Abschlusseigenschaften

In der Mathematik, insbesondere der Algebra, versteht man unter Abgeschlossenheit einer Menge bezüglich einer Verknüpfung, dass die Verknüpfung beliebiger Elemente dieser Menge wieder ein Element der Menge ergibt. Beispielsweise ist die Menge der ganzen Zahlen abgeschlossen bezüglich der Addition, Subtraktion und Multiplikation, aber nicht bezüglich der Division.[7]

$L_1 \cup L_2$ (**Vereinigung**): Die Vereinigung $L = L_1 \cup L_2$ zweier Sprachen L_1 und L_2.[8]

$L_1 \cap L_2$ (**Schnitt**): Der Schnitt $L = L_1 \cap L_2$ zweier Sprachen L_1 und L_2.[9]

$\bar{L}$ (**Komplement**): Das Komplement $\bar{L} = \Sigma^* \setminus L$ einer Sprache L.[10]

$L_1 \circ L_2$ (**Produkt / Konkatenation**): Das Produkt $\{ uv \mid u \in L_1 \vee v \in L_2 \}$ zweier Sprachen L_1 und L_2.[11]

L^* (**Kleene-Stern**): Der Kleene-Stern L^* einer Sprache L, d. h. die beliebig häufige Konkatenation von Wörter aus der Sprache L vereinigt mit dem leeren Wort.[12][13]

	$L_1 \cup L_2$	$L_1 \cap L_2$	$\bar{L}$	$L_1 \circ L_2$	L^*
Typ 3	ja	ja	ja	ja	ja
Typ 2	ja	nein	nein	ja	ja
Typ 1	ja	ja	ja	ja	ja
Typ 0	ja	ja	nein	ja	ja

[14]

[4] *Theoretische Informatik – Reguläre Sprachen*, Seite 15.
[5] *Wikipedia-Artikel „Terminalsymbol“*.
[6] *Wikipedia-Artikel „Produktionsregel“*.
[7] *Wikipedia-Artikel „Abgeschlossenheit (algebraische Struktur)“*.
[8] *Wikipedia-Artikel „Menge (Mathematik)“*.
[9] *Wikipedia-Artikel „Menge (Mathematik)“*.
[10] *Wikipedia-Artikel „Komplement (Mengenlehre)“*.
[11] *Wikipedia-Artikel „Formale Sprache“*.
[12] *Theoretische Informatik – Reguläre Sprachen*, Seite 68.
[13] *Wikipedia-Artikel „Kleenesche und positive Hülle“*.
[14] *Theoretische Informatik – Reguläre Sprachen*, Seite 33.

Entscheidbarkeitseigenschaften[15]

$\omega \in L(G)$ (**Wortproblem**): Als Wortproblem einer formalen Sprache bezeichnet man das Entscheidungsproblem, zu einem *gegebenen Wort* festzustellen, ob dieses *zur Sprache gehört*, oder nicht. Das Wortproblem einer Sprache L ist entscheidbar, wenn es einen Algorithmus gibt, der in endlicher Zeit herausfindet, ob $\omega \in L(G)$ ist oder nicht.[16]

gegebenen Wort
zur Sprache gehört

$L(G) = \emptyset$ (**Leerheitsproblem**): Als Leerheitsproblem bezeichnet man das Problem, zu entscheiden, ob eine in Form einer formalen Grammatik gegebene *formale Sprache L leer ist*, also $L(G) = \emptyset$. Das Problem ist es also herauszufinden, ob es Wörter gibt, die den Regeln der Grammatik genügen, oder nicht.[17]

formale Sprache L leer ist

$L(G_1) = L(G_2)$ (**Äquivalenzproblem**): Als Äquivalenzproblem bezeichnet man das Problem, zu entscheiden, ob zwei formale Definitionen von zwei Sprachen L_1 und L_2 *äquivalent* sind, also $L(G_1) = L(G_2)$ gilt.[18][19]

äquivalent

$L(G_1) \cap L(G_2) = \emptyset$ (**Schnittproblem**): Als Schnittproblem bezeichnet man die Frage, ob die Schnittmenge zweier formaler Sprachen, die durch Grammatiken gegeben sein sollen, leer ist.[20]

$|L(G)| < \infty$ (**Endlichkeitsproblem**): Als Endlichkeitsproblem einer formalen Sprache L bezeichnet man das Problem, zu entscheiden, ob die Sprache endlich ist. Eine formale Sprache wird als endlich bezeichnet, wenn die Menge ihrer „Wörter“ endlich ist, man schreibt dann auch $|L| < \infty$.[21]

	$\omega \in L(G)$	$L(G) = \emptyset$	$L(G_1) = L(G_2)$	$L(G_1) \cap L(G_2) = \emptyset$	$\|L(G)\| < \infty$
Typ 3	ja	ja	ja	ja	ja
Typ 2	ja	ja	nein	nein	ja
Typ 1	ja	nein	nein	nein	nein
Typ 0	nein	nein	nein	nein	nein

Formale Sprachen und Maschinenmodelle[22]

Sprache	Maschinenmodell	Speichergröße
Typ 3	DFA, NFA	endlich
Typ 2	DPDA, NPDA	Stapel
Typ 1	LBA	proportional zu Eingabe
Typ 0	DTM, NTM	unendlich

15 Schneider, *Taschenbuch der Informatik*, Seite 591 Kapitel 19.1.3.3.
16 *Wikipedia-Artikel „Wortproblem“*.
17 *Wikipedia-Artikel „Leerheitsproblem“*.
18 *Theoretische Informatik – Reguläre Sprachen*, Seite 70-71.
19 *Wikipedia-Artikel „Äquivalenzproblem“*.
20 *Wikipedia-Artikel „Schnittproblem“*.
21 *Wikipedia-Artikel „Endlichkeitsproblem“*.
22 *Theoretische Informatik – Typ-1- und Typ-0-Sprachen*, Seite 32.

Englisch	Bedeutung	Deutsch	Bedeutung
DFA[23]	Deterministic Finite Automaton	DEA	Deterministischer Endlicher Automat
DPDA[24]	Deterministic Push-Down Automaton	DKA	Deterministischer Kellerautomat
TM, DTM[25]	Deterministic Turing machine		Deterministische Turing-Maschine
LBA[26]	Linear Bounded Automaton		Linear beschränkte Turingmaschine
NFA[27]	Deterministic Finite Automaton	NEA	Nichtdeterministischer endlicher Automat
PDA, NPDA[28]	(Nondeterministic) PushDown Automaton	KA, NKA	Nichtdeterministischer Kellerautomat
NTM, NDTM[29]	NonDeterministic Turing machine	-	Nichtdeterministische Turingmaschine

Literatur

[1] Dirk W. Hoffmann. *Theoretische Informatik*. 2018.

[2] Uwe Schneider. *Taschenbuch der Informatik*. 7. Aufl. Hanser, 2012. ISBN: 9783446426382.

[3] *Theoretische Informatik – Reguläre Sprachen*.

[4] *Theoretische Informatik – Typ-1- und Typ-0-Sprachen*.

[5] *Wikipedia-Artikel „Abgeschlossenheit (algebraische Struktur)“*. https://de.wikipedia.org/wiki/Abgeschlossenheit_(algebraische_Struktur).

[6] *Wikipedia-Artikel „Äquivalenzproblem“*. https://de.wikipedia.org/wiki/Äquivalenzproblem.

[7] *Wikipedia-Artikel „Chomsky-Hierarchie“*. https://de.wikipedia.org/wiki/Chomsky-Hierarchie.

[8] *Wikipedia-Artikel „Deterministic pushdown automaton“*. https://de.wikipedia.org/wiki/Deterministic_pushdown_automaton.

[9] *Wikipedia-Artikel „Deterministischer endlicher Automat“*. https://de.wikipedia.org/wiki/Deterministischer_endlicher_Automat.

[10] *Wikipedia-Artikel „Endlichkeitsproblem“*. https://de.wikipedia.org/wiki/Endlichkeitsproblem.

[23] *Wikipedia-Artikel „Deterministischer endlicher Automat“*.
[24] *Wikipedia-Artikel „Deterministic pushdown automaton“*.
[25] *Wikipedia-Artikel „Turingmaschine“*.
[26] *Wikipedia-Artikel „Linear beschränkte Turingmaschine“*.
[27] *Wikipedia-Artikel „Nichtdeterministischer endlicher Automat“*.
[28] *Wikipedia-Artikel „Kellerautomat“*.
[29] *Wikipedia-Artikel „Nichtdeterministische Turingmaschine“*.

[11] *Wikipedia-Artikel „Formale Sprache“*. https://de.wikipedia.org/wiki/Formale_Sprache.

[12] *Wikipedia-Artikel „Kellerautomat“*. https://de.wikipedia.org/wiki/Kellerautomat.

[13] *Wikipedia-Artikel „Kleenesche und positive Hülle“*. https://de.wikipedia.org/wiki/Kleenesche_und_positive_Hülle.

[14] *Wikipedia-Artikel „Komplement (Mengenlehre)“*. https://de.wikipedia.org/wiki/Komplement_(Mengenlehre).

[15] *Wikipedia-Artikel „Leerheitsproblem“*. https://de.wikipedia.org/wiki/Leerheitsproblem.

[16] *Wikipedia-Artikel „Linear beschränkte Turingmaschine“*. https://de.wikipedia.org/wiki/Linear_beschränkte_Turingmaschine.

[17] *Wikipedia-Artikel „Menge (Mathematik)“*. https://de.wikipedia.org/wiki/Menge_(Mathematik).

[18] *Wikipedia-Artikel „Nichtdeterministische Turingmaschine“*. https://de.wikipedia.org/wiki/Nichtdeterministische_Turingmaschine.

[19] *Wikipedia-Artikel „Nichtdeterministischer endlicher Automat“*. https://de.wikipedia.org/wiki/Nichtdeterministischer_endlicher_Automat.

[20] *Wikipedia-Artikel „Produktionsregel“*. https://de.wikipedia.org/wiki/Produktionsregel.

[21] *Wikipedia-Artikel „Schnittproblem“*. https://de.wikipedia.org/wiki/Schnittproblem.

[22] *Wikipedia-Artikel „Terminalsymbol“*. https://de.wikipedia.org/wiki/Terminalsymbol.

[23] *Wikipedia-Artikel „Turingmaschine“*. https://de.wikipedia.org/wiki/Turingmaschine.

[24] *Wikipedia-Artikel „Wortproblem“*. https://de.wikipedia.org/wiki/Wortproblem.