Typ-0 und Typ-1 Grammatiken

Jede Grammatik ist eine Typ 0 Grammatik. Allerdings kann es nur abzählbar viele Grammatiken (über endlichem Alphabet) geben.

Zum Vergleich: In Einheit 1 haben wir festgestellt, dass es überabzählbar viele formale Sprachen über festem Alphabet gibt.

Typ 1 Grammatiken werden auch als *nichtverkürzend* bezeichnet. Durch unsere Definition erklärt sich dieser Name von selbst. Man spricht auch von kontextsensitiven Grammatiken – obwohl dieser Begriff von einer anderen (gleichwertigen) Definition der Typ 1 Grammatiken herrührt. (Bitte informieren Sie sich hierüber.)

Beachte: Wenn alle Regeln nichtverkürzend sind und wir mit einer Satzform der Länge 1 beginnen (S), können wir nie eine Satzform der Länge 0 erhalten. Wir klären das später (ε-Sonderregel).

Typ-2 Grammatiken

Grammatiken vom Typ-2 zeichnen sich dadurch aus, dass in den Regeln (aus P) die linken Seiten immer die Länge 1 haben.
Und nicht nur das: Das eine Symbol muss immer aus V kommen.

Da Typ-2 Grammatiken insbesondere auch Typ-1 Grammatiken sind (lt. Definition), sind die rechten Seiten *nichtleere* Satzformen.
Aber sonst unterliegen sie keinerlei Einschränkungen.

Also besagen die Regeln, dass ein bestimmtes *Nichtterminalsymbol* (d.h. eine Variable) durch eine bestimmte nichtleere Satzform ersetzt werden darf, unabhängig davon, in welchem Zusammenhang diese Variable gerade erscheint.

Daher der Name kontextfreie Grammatik.

Typ-3 Grammatiken

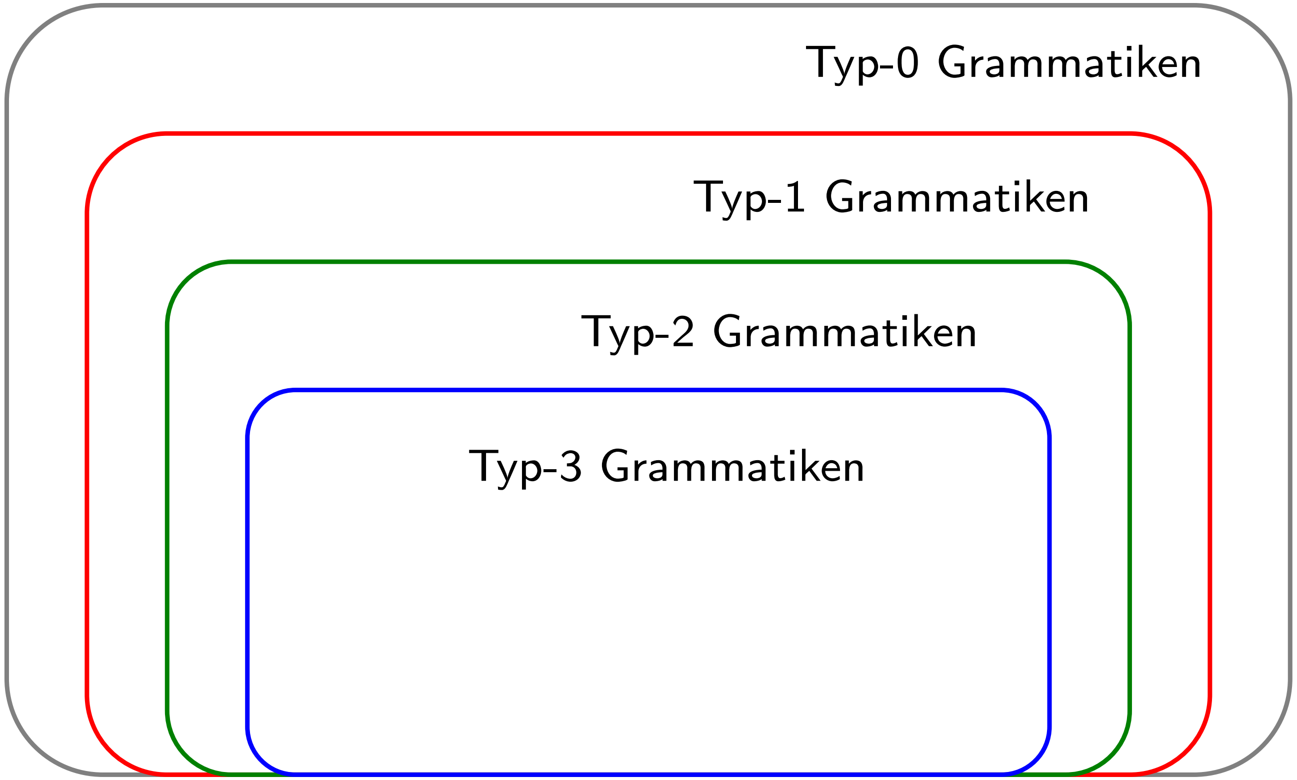
Typ-3 Grammatiken sind wiederum ein Spezialfall der Typ-2 Grammatiken. Jetzt muss die rechte Seite jeder Regel aus genau einem Terminalsymbol (Buchstaben aus dem Alphabet) oder einem Terminalsymbol gefolgt von einer Variablen bestehen.

Hieraus folgt sofort, dass die Satzformen, die aus dem Startsymbol durch Schritte gemäß der Übergangsrelation $\Rightarrow_G$ entstehen, aus einer Folge von Buchstaben (aus Σ) gefolgt von *höchstens einer* Variablen bestehen.

Und sobald diese Variable verschwindet (d.h. es wurde eine Regel verwendet, bei der die rechte Seite nur aus einem Buchstaben besteht), kann dieser *Ableitungsprozess* nicht fortgeführt werden.

Schaubild

Wir veranschaulichen die Chomsky-Hierarchie so:



Chomsky-Hierarchie für Sprachklassen

Wir haben definiert, wie eine Grammatik eine Sprache beschreibt.

Daher kann man auf Grund der Klasseneinteilung der Grammatiken auch eine entsprechende Klasseneinteilung der Sprachen einführen:

> **Definition:** Für $i \in \{0, 1, 2, 3\}$ und eine formale Sprache $L \subseteq \Sigma^*$ gilt: L ist vom Typ i, wenn es eine Grammatik G vom Typ i gibt mit $L = L(G)$.

Da jede Grammatik vom Typ $i > 0$ auch eine Grammatik vom Typ $i - 1$ ist, erhalten wir also auch hier eine Hierarchie!

Jede Typ-3 Sprache ist auch eine Typ-2 Sprache. Jede Typ-2 Sprache ist auch eine Typ-1 Sprache; und jede Typ-1 Sprache ist auch eine Typ-0 Sprache.

Die Typ-0 Sprachen

Da alle Grammatiken vom Typ-0 sind, könnte man glauben, dass auch alle Sprachen automatisch immer Typ-0 Sprachen wären.

Das trifft aber keinesfalls zu !!

Schon aus Anzahlgründen ist das nicht möglich, denn es gibt ja überabzählbar viele (formale) Sprachen über einem Alphabet Σ. Aber es gibt nur abzählbar viele Grammatiken über dem selben Alphabet. Und es ist offensichtlich, dass eine Grammatik nur genau eine formale Sprache beschreibt.

Also gibt es höchstens abzählbar viele Typ-0 Sprachen!

Typ-1 Sprachen

Die Beispielgrammatik für die Sprache

$$L = \{a^n b^n c^n \mid n \geq 1\}$$

hat die folgende Regelmenge:

$$\begin{aligned} P \;=\; \{&(S, aSBC),\ (S, aBC),\ (CB, BC), \\ &(aB, ab),\ (bB, bb),\ (bC, bc), (cC, cc)\} \end{aligned}$$

die offensichtlich nichtverkürzend ist. Also ist L eine Sprache vom Typ-1 und natürlich auch vom Typ-0.

Es gibt aber auch Sprachen vom Typ-0, die nicht vom Typ-1 sind. Ein bekanntes Beispiel hierfür ist die Menge H, das sogenannte *Halteproblem*, das wir später noch kennenlernen werden.

Typ-2 und Typ-3 Sprachen

In Einheit 1 haben wir mehrere Beispiele für Typ-2 Grammatiken angegeben: Die Grammatik für die korrekt geklammerten arithmetischen Ausdrücke, die Grammatik für die Palindrome gerader Länge, sowie die Grammatik für die Sprache $\{a^n b^n \mid n \geq 1\}$.

Also sind alle diese Sprachen Typ-2 Sprachen.

Und damit sind sie alle auch Typ-1 Sprachen und ebenso Typ-0 Sprachen.

Allerdings sind alle diese Sprachen *nicht* Typ-3.

Das können wir im Moment aber noch nicht nachweisen – Beweise folgen später...

Eine typische Typ-3 Sprache ist $L = \{(ab)^n \mid n \geq 1\}$.

Typ-3 Grammatik für diese Sprache: $G = (\{S, B\}, \{a, b\}, P, S)$ mit $P = \{(S, aB), (B, b), (B, bS)\}$.

Was wissen wir über $G = (\{S\}, \{a, b\}, P, S)$ mit $P = \{(S, abS), (S, ab)\}$?

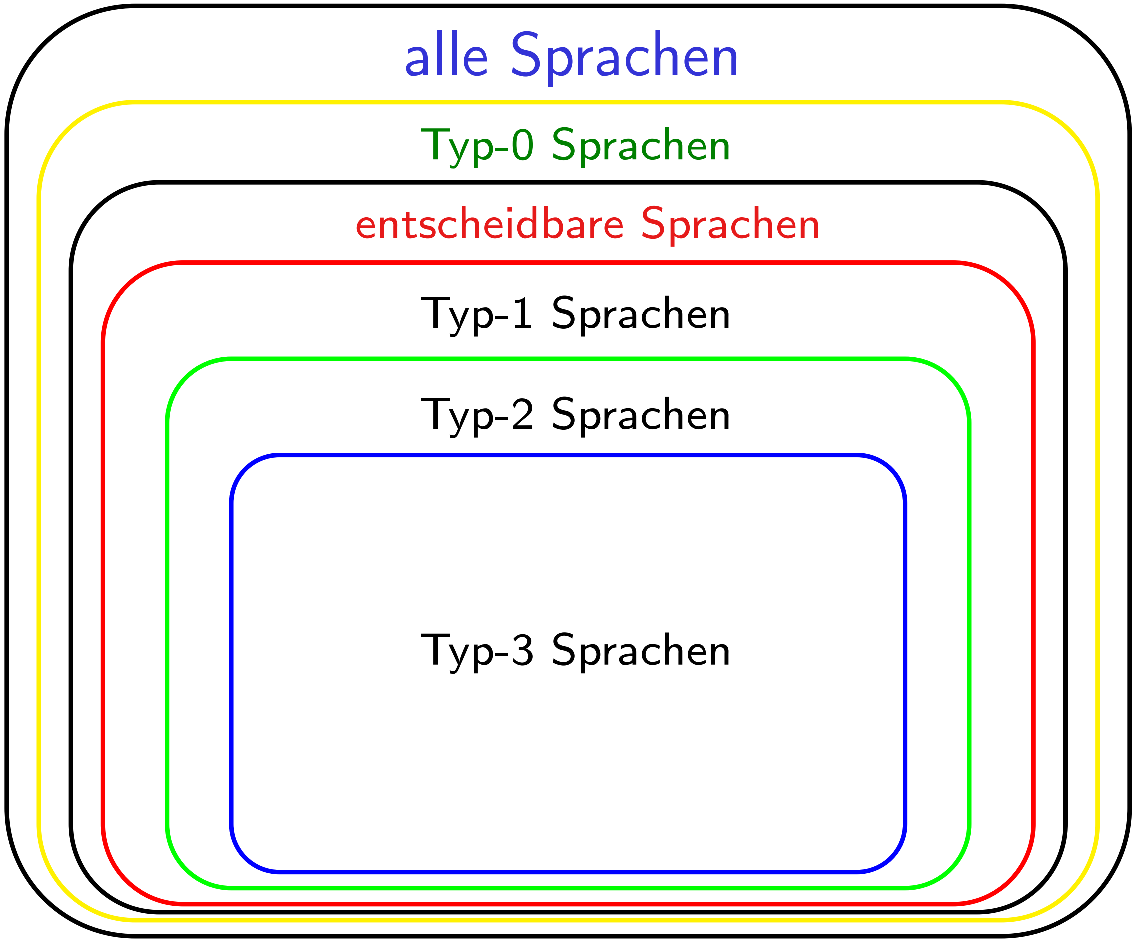
Testfragen

- Gibt es Grammatiken, die Typ-2 sind, aber nicht Typ-1?
- Gibt es Sprachen, die Typ-2 sind, aber nicht Typ-1?
- Gibt es Grammatiken, die Typ-1 sind, aber nicht Typ-2?
- Gibt es Sprachen, die Typ-1 sind, aber nicht Typ-2?
- Ist die folgende Grammatik Typ-0, Typ-1, Typ-2, Typ-3?

  $G = (\{S, A\}, \{a\}, P, S)$ mit $P = \{(S, SAA), (S, AA), (AA, aa)\}$
- Was ist die zugehörige Sprache?
- Ist diese Sprache Typ-0, Typ-1, Typ-2, Typ-3?
- Wie sieht eine „optimale“ Grammatik für diese Sprache aus?

Merkbild

Die Situation, die durch die Chomsky-Hierarchie auf Sprachebene entsteht, wollen wir graphisch darstellen:



Beachte: Mathematisch gesehen ist nur ein sehr kleiner Teil aller Sprachen durch Grammatiken zu definieren!

Was die *entscheidbaren* Sprachen sind, werden wir in Einheit 5 sehen.