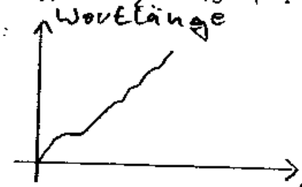
Blatt 1 - Grammatiken

Grammatik $G=(V,\Sigma,P,S)$.
Welchen Typs sie ist, hängt von den Regeln (Produktionen) ab.

Typ 0: Jede Grammatik ist vom Typ 0, da gerade das Existieren der Grammatik eine Sprache zu einer Typ 0-Sprache macht.
Insbesondere existieren aber auch Sprachen, für die man keine Grammatik angeben kann; diese sind dann nicht einmal Typ 0.
(Als Beispiel kann man sich überlegen, dass es ab zweielementigem Terminalalphabet möglich ist beliebig lange Wörter zu erzeugen, die in der Sprache liegen sollen, aber keine regelmäßige Entstehung haben.
$\Sigma=\{a\} \Rightarrow$ beliebig lange Wörter aber eben nur a*
$\Sigma=\{a,b\} \Rightarrow$ ~ , das b z.B. kann irgendwo in den a's vorkommen ohne erkennbare Regelmäßigkeit.
(* Erzeuge ich über {a} nur Wörter bestimmter Länge ohne irgendeine Regelmäßigkeit, so lässt sich dies bei unendlich vielen Wörtern auch nicht in Regeln verpacken...)

Typ 1: $\forall$ Regeln $w_1 \to w_2$ aus P: $|w_1| \leq |w_2|$ z.B. $aBc \to abDc$

anschaulich:



„Die Kurve steigt monoton, d.h. die Wörter werden immer länger, nicht kürzer."

Damit wird der Test, ob ein Wort in der Sprache liegt erheblich vereinfacht, da wenn ich für ein Wort bestimmter Länge teste und alle Ableitungsmöglichkeiten für Formen dieser Länge getestet habe und es nicht dabei war, ich sicher sagen kann, dass es nicht in der Sprache liegt; denn ab dem Punkt, an dem ich mich befinde, werden die abgeleiteten Formen nur noch länger...

Typ 1 $\hat{=}$ Kontextsensitiv (Im unserer Beispielregel oben $aBc \to abDc$ kann das B eben nur im Kontext von a und c in bD verwandelt werden)

Typ 2: $\forall\ w_1 \to w_2 \in P$ $|w_1|=1 \wedge w_1 \in V$ „w_1 ist einzelne Variable"
z.B. $A \to a$, $A \to MNO$
Typ 2 $\hat{=}$ Kontextfrei (Im Beispiel kann A jetzt unabhängig vom Kontext immer durch a oder MNO ersetzt werden)

Typ 3: $\forall\ w_1 \to w_2 \in P$ $w_2 \in \Sigma \vee w_2 \in \{aA \mid a \in \Sigma, A \in V\}$
„w_2 ist einzelnes Terminalsymbol oder Terminalsymbol gefolgt von einer Variablen"
z.B. $A \to a$, $A \to aM$
Typ 3 $\hat{=}$ regulär

Es gilt:

$\boxed{\text{Alle Sprachen} \supsetneq \text{Typ 0-Spr.} \supsetneq \text{Typ 1-Spr.} \supsetneq \text{Typ 2-Spr} \supsetneq \text{Typ 3-Spr}}$

Eine Sprache höheren Typs ist also echte Teilmenge ($\subsetneq$) derjenigen niedrigeren Typs.
Für die Regeln auf der vorherigen Seite bedeutet dies, dass jedes mal eine hinzukommt; Für Typ 3 gelten also alle Einschränkungen!

Für kontext freie Sprachen (also Typ 2 und Typ 3) werden zusätzlich noch Regeln der Form $A \rightarrow \varepsilon$ erlaubt, die wegen $|w_1| > |w_2|$ ja eigentlich verboten wären.
Dass dies aber nur „syntaktischer Zucker" ist, sieht man leicht an folgendem Beispiel:

$M \rightarrow dAb$ (Auszug einer Typ 2-Grammatik, der alle
$N \rightarrow Abc$ Regeln enthält, die A beinhalten)
$A \rightarrow a$
$A \rightarrow \varepsilon$

$\boxed{A \rightarrow \varepsilon}$

$M \rightarrow dAb$
$M \rightarrow db$
$N \rightarrow Abc$
$N \rightarrow bc$
$A \rightarrow a$

Wie wir sehen, ist es also durch „rein mechanisches" Umformen möglich, eine äquivalente Grammatik ohne die Regel $A \rightarrow \varepsilon$ zu erzeugen und trotzdem gleichen Typs zu bleiben (immer noch Typ 2)

Nach den Definitionen der letzten Seite wäre $S \rightarrow \varepsilon$ nicht erlaubt, das leere Wort ε also niemals Element einer Typ 1, 2, 3 - Sprache.
Das will man aber so nicht unbedingt haben, also erlaubt man die Regel $S \rightarrow \varepsilon$ als Ausnahme zur Definition.
„S" darf dann aber nicht mehr auf der rechten Seite einer Regel vorkommen, was nicht schlimm ist, da man Folgendes tun kann:

1. $S \rightarrow$ rechte Seiten der S-Regeln mit S ersetzt durch S'
2. ersetze in allen anderen Regeln S durch S'
3. Füge $S \rightarrow \varepsilon$ hinzu