1. Formale Sprachen

1.4 Grenzen endlicher Automaten*

Die regulären Sprachen sind genau diejenigen, die von endlichen deterministischen Automaten erkannt werden.

Die Produktionsregeln der Grammatik lassen sich alle in der Form
Z1 = ‘t‘ Z2   oder   Z1 = ‘t‘
darstellen. t ist ein Terminalzeichen, Z1 und Z2 sind Nichtterminale.

Eine Grammatik mit diesem Aufbau nennt man **rechtslinear**.
Um andere formale Sprachen zu beschreiben, benötigt man andere Produktionsregeln und natürlich auch andere Automaten.

*kein verpflichtender Lerninhalt für das Abitur

Beispiel:

Die Sprache der Palindrome über dem Alphabet {a,b}.
Wörter der Sprache sind z.B. abba, abbbba, babbab, aaa, ...

Produktionsregeln:
P = a P a | b P b | a | b | aa | bb

Eine Grammatik dieser Art nennt man **kontextfrei** , d.h. alle Regeln haben gemeinsam, dass auf der linken Seite nur ein Nichtterminal steht.

Diese Sprache kann nicht durch einen deterministischen endlichen Automaten erkannt werden und ist deshalb nicht regulär.

In der theoretischen Informatik gibt es Methoden, mit denen man dies beweisen kann.
Ein Beweis ist hier notwendig, denn es könnte ja möglich sein, die Produktionsregeln äquivalent umzuformen, so dass sie rechtslinear sind.

Ein anderes Beispiel für eine kontextfreie Sprache ist die Sprache der **Terme mit beliebig verschachtelten Klammern**.

Man benötigt dafür u.a. Regeln der Form:

Teilausdruck = Variable | Zahl | "(" Ausdruck ")"
Ausdruck = Teilausdruck { (“+“ | “-“| “·“| “:“) Teilausdruck }

Beispiel:
$\left(a - 2 \cdot (a + b)\right)$

Es gibt auch **kontextsensitive** Grammatiken, deren Regeln dann auch Formen wie

$$x\,A\,y\ = x\,z\,y$$

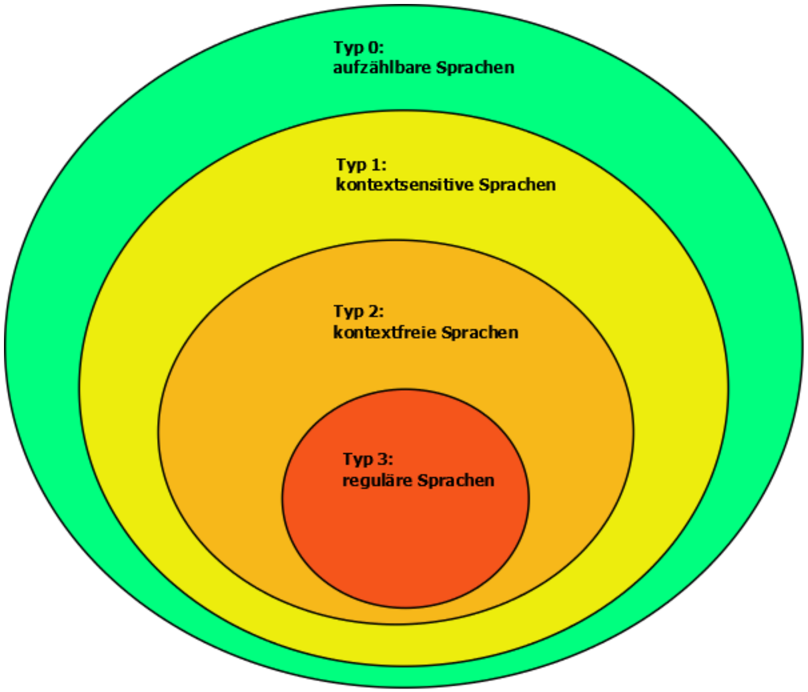
haben können.

x und y sind dabei beliebige Zeichenketten aus Terminal- und Nichtterminalsymbolen einschließlich des leeren Worts ε.

z steht ebenfalls für beliebige Zeichenketten aus Terminal- und Nichtterminalsymbolen ausschließlich des leeren Worts ε. Die Regel $\mathrm{A} = \varepsilon$ ist als Sonderfall zugelassen.

Beispiel für eine kontextsensitive Grammatik:
Übung, Aufgabe 4

Die Chomsky - Hierarchie
Von dem amerikanischen Sprachwissenschaftler Noam Chomsky (*1928) stammt folgende Einteilung der Sprachen.



Die Graphik gibt dabei eine Teilmengenbeziehung wieder.
Die Menge der regulären Sprachen ist echt enthalten in der Menge der kontextfreien Sprachen.

Auch für die nicht regulären Sprachen gibt es Automaten, welche die Sprache erkennen.
Im Falle der Typ 0 - Sprachen sind dies die **Turingmaschinen**, benannt nach dem Mathematiker Alan Turing (1912 - 1954).
Auf den Seiten von MathePrisma der Uni Wuppertal gibt es eine gute Simulation und Erklärung:
http://www.matheprisma.uni-wuppertal.de/Module/Turing/
Oder:
https://turingmachinesimulator.com

Übung

1. Gib eine Grammatik für die Sprache $L = \{ a^n b^n | n \in \mathbb{N}\}$ an.
   Wörter dieser Sprache sind z. B. $ab, aabb, aaabbb, \ldots$

2. Welche Sprache über $\{a, b, c\}$ wird durch folgende Regeln beschrieben?
   Start = S
   $S = A\,B$
   $A = a\,b \mid a\,A\,b$
   $B = c \mid c\,B$

3. Welche Sprache über $\{a, b\}$ wird durch folgende Regeln beschrieben?
   Start = S
   $S = XSY|\varepsilon \qquad XY = YX \qquad X = a \qquad Y = b$

4. Welche Sprache über $\{a, b, c\}$ wird durch folgende Regeln beschrieben?
   Start = S
   $S = a\,S\,B\,c \mid a\,b\,c$
   $c\,B = B\,c$
   $b\,B = b\,b$