

Automaten und formale Sprachen

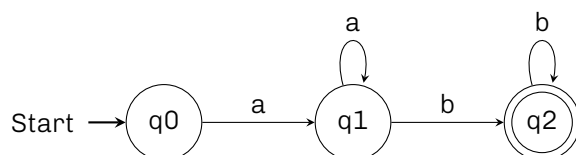
Kontextfreie Sprachen

Notation von Sprachen

Sprachen sind Mengen von Wörtern. Um eine Sprache zu beschreiben gibt es verschiedene Notationen:

- Umgangssprachlich: „Alle Wörter, die mit beliebig vielen a anfangen, gefolgt von genauso vielen b.“
- Als spezifische Menge von Wörter: $L = \{000, 001, 010, 011, 100, 101, 110, 111\}$
- Als allgemeine Menge von Wörtern: $L = \{w \mid w \text{ ist ein Palindrom}\}$
- Als regulärer Ausdruck: $L = a^*(c|b)a^*$
- Als erweiterter regulärer Ausdruck: $L = a^n b^m, \quad n, m \geq 1$

Eine Sprache ist *regulär*, wenn sie durch einen (deterministischen) endlichen Automaten akzeptiert, bzw. durch eine (rechts)reguläre Grammatik erzeugt wird. Beispielsweise $L = a^n b^m, n, m \geq 1$:



$$\begin{aligned}
 S &\rightarrow aA \\
 A &\rightarrow aA \mid bB \\
 B &\rightarrow bB \mid \varepsilon
 \end{aligned}$$

Aufgabe 1

Erstelle Grammatiken und endliche Automaten zu folgenden Sprachen:

- $L_1 = a^n b^m, \quad n \geq 0, m \geq 2$
- $L_2 = a^n b^n, \quad 1 \leq n \leq 2$
- $L_3 = a^n b^n, \quad n \geq 1$

Aufgabe 2

Tatsächlich gibt es keine reguläre Grammatik, die die Sprache L_3 aus Aufgabe 1 erzeugt.

Beschreibt möglichst präzise das Problem, warum eine reguläre Grammatik nicht ausreicht und diskutiert Lösungsansätze. Wie könnte eine Grammatik zu L_3 aussehen, wenn die Produktionen keinen Einschränkungen unterliegen?

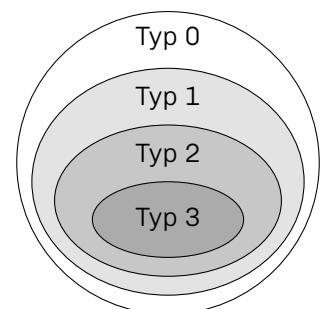
Die Sprache L_3 ist ein Beispiel für eine nicht-reguläre Sprache. Formale Sprachen lassen sich in eine Hierarchie mit vier Ebenen einsortieren. Jede Ebene ist Teil der darüber liegenden (eine Typ 2 Sprache ist auch eine Typ 3 Sprache).

Typ 0 *Rekursiv aufzählbare Sprachen*: Sprachen ohne Einschränkungen. Sie sind äquivalent zu *Turingmaschinen*.

Typ 1 *Kontextsensitive Sprachen*: Sprachen, deren Produktionen vom „Kontext“ abhängen können, in dem sich ein Nichtterminal befindet.

Typ 2 *Kontextfreie Sprachen*: Sprachen, deren Produktionen keine Einschränkungen haben, aber vom „Kontext“ unabhängig sind.

Typ 3 *Reguläre Sprachen*: Sprachen, die zusätzlichen Einschränkungen unterliegen.



Darstellung der Chomsky-Hierarchie