Automaten und formale Sprachen

Kellerautomaten

Reguläre Sprachen haben ihre Grenzen. Die Sprache $a^n b^n, n \geq 0$ kann nicht durch einen endlichen Automaten akzeptiert werden und wir können auch keine reguläre Grammatik konstruieren. Die Sprache ist also nicht regulär, sondern *kontextfrei*. Um einen Automaten zu konstruieren, der eine kontextfreie Sprache akzeptiert, müssen wir das Automatenmodell erweitern.

Das Grundproblem: *Ein endlicher Automat kann nicht zählen!* Er kann sich nicht merken, wie viele a schon gekommen sind, um entsprechend viele b abzuzählen. Wir brauchen also eine Struktur, die uns das Zählen erlaubt und diese finden wir im *Keller*.

Ein *nichtdeterministischer Kellerautomat* (NKA) wird als 7-Tupel $A = (Q, \Sigma, K, \delta, q_0, \#, F)$ definiert:

- Q ist die (endliche) Menge von Zuständen des Automaten.
- Σ ist das Eingabealphabet.
- K ist das Kelleralphabet.
- $\delta : Q \times (\Sigma \cup \{\varepsilon\}) \times K \to \mathcal{P}(Q \times K^*)$ ist die Übergangsfunktion.
- $q_0 \in Q$ ist der Startzustand.
- $\# \in K$ ist das Startsymbol des Kellers.
- $F \subset Q$ ist die Menge der Endzustände.

Schematisch kann man sich einen Kellerautomaten wie in Abbildung 1 vorstellen.

Jeder Übergang liest nun nicht mehr nur einen Buchstaben aus dem Eingabealphabet, sondern auch einen Buchstaben aus dem Kelleralphabet: Den obersten Buchstaben auf dem Keller. Das Kellerstartsymbol $\#$ liegt zu Beginn im Keller. Wenn es gelesen wird, dann gilt der Keller als leer.

Zur Darstellung des Übergangsgraphen erweitern wir die Darstellung eine endlichen Automaten um das Kelleralphabet:

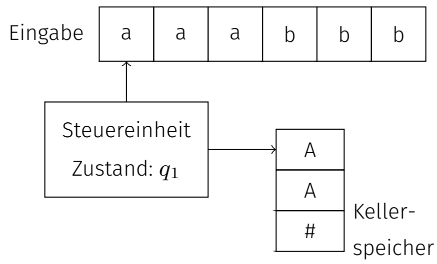


Abbildung 1: Schematische Darstellung eines Kellerautomaten

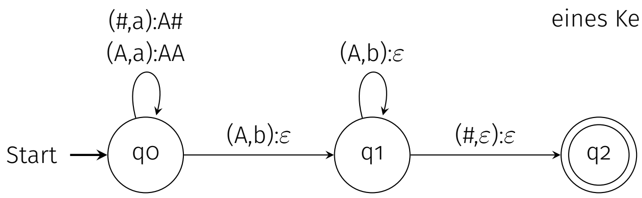


Abbildung 2: Kellerautomat zur Sprache $L_K = a^n b^n$ (Klammersprache).

Ein Übergang hat nun die Form $(X, Y) : Z^*$, wobei $X \in K$, $Y \in \Sigma$ und $Z \in K$ sind. Zu beachten ist, dass die Folge von Kellersymbolen Z^* *von rechts nach links* auf den Keller gelegt werden. Für $Z = \varepsilon$ wird nichts auf den Keller gelegt. Da immer das oberste Kellersymbol gelesen wird, wird also ein Buchstabe vom Keller entfernt!

Aufgabe 1

Baue den Kellerautomaten zur Klammersprache in Abbildung 2 in FLACI nach und teste ihn.

Aufgabe 2

In der Regel ist nicht nur die Verschachtelung von Klammerpaaren ineinander erlaubt, sondern auch mehrere verschachtelte Klammerpaare hintereinander.

a) Erweitere den Klammerautomaten aus Aufgabe 1 zur Sprache

$$L_{K2} = (L_K)^* = (a^{n_i} b^{n_i})^m, n_i \geq 1, m \geq 0, i = 1,\ldots,m$$

Gültige Worte sind zum Beispiel: $aaabbbaabb$, $ababab$, $aabbaabb$.

b) Erweitere den NKA L_{K2} zur Sprache L_{K3}, die beliebige gültige Klammerausdrücke erkennt. Im Gegensatz zu L_{K2} sind dann auch Worte wie $aababb$ oder $aaabaababbabbb$ gültig. (Es dürfen also neue Klammerpaare beginnen, bevor alle Klammern geschlossen wurden.)

Aufgabe 3

Entwickle in FLACI einen NKA für die Sprachen L_4 bis L_6:

$$\begin{aligned} L_4 &= a^n (a|b|c)^n, n \geq 1 \\ L_5 &= \{w_1 a w_2 | w_1 \text{ und } w_2 \text{ sind gleichlange Worte über } \{1,0\}\} \\ L_6 &= a^n b^m, 0 \leq a < b \end{aligned}$$

Gib auch jeweils die formale Definition der Automaten als 7-Tupel an.

Aufgabe 4

Erstelle zu den Sprachen L_4, L_5 und L_6 aus Aufgabe 3 jeweils eine kontextfreie Grammatik, die die Sprache erzeugt.

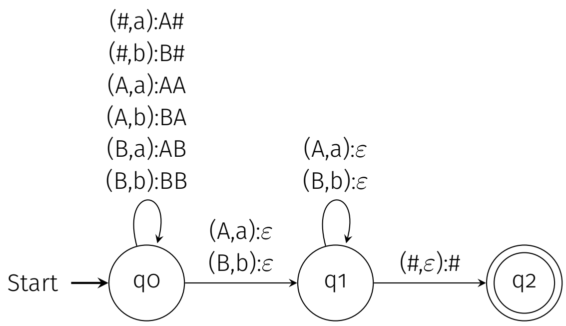
Aufgabe 5

$G_{RE} = (N,T,S,P)$ ist die rechtsreguläre Grammatik für Rechenterme vom Arbeitsblatt 5, mit den in Abbildung 3.a gezeigten Produktionen (zur Vereinfachung wird die Kurzschreibweise $0..3B$ für $0B|1B|2B|3B$ genutzt).

Erweitere G_{RE} zu einer kontextfreien Grammatik, die auch Terme mit korrekten Klammerausdrücken erlaubt.

$$\begin{aligned} &S \rightarrow 0A \,|\, 1..9B \,|\, \varepsilon \\ &A \rightarrow .C \,|\, +E \,|\, -E \,|\, \cdot E \,|\, :E \,|\, \varepsilon \\ &B \rightarrow .C \,|\, 0..9B \,|\, +E \,|\, -E \,|\, \cdot E \,|\, :E \,|\, \varepsilon \\ &C \rightarrow 0..9D \\ &D \rightarrow 0..9D \,|\, +E \,|\, -E \,|\, \cdot E \,|\, :E \,|\, \varepsilon \\ &E \rightarrow 0A \,|\, 1..9B \end{aligned}$$

3.a: Produktionen von G_{RE}.



3.b: Übergangsgraph eines NKA.

Aufgabe 6

Welche Sprache beschreibt der Kellerautomat zum in Abbildung 3.b gezeigten Übergangsgraphen?