

Theoretische Informatik

Teil 5

Kellerautomaten

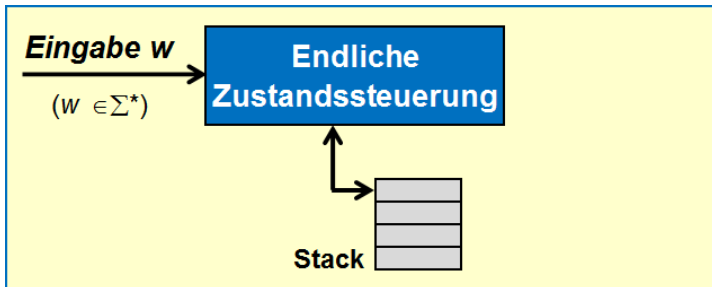
Frühlingssemester 2019

L. Di Caro

D. Flumini

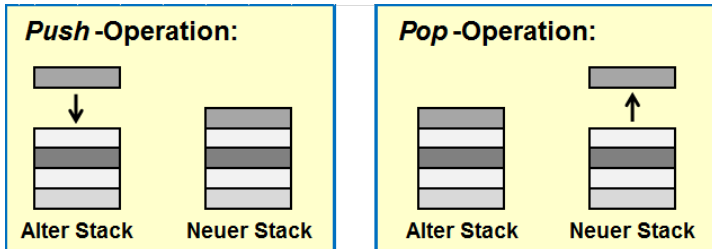
O. Stern

- Automatenmodell für die Erkennung von kontextfreien Sprachen.
- Endlicher Automat mit zusätzlichem (unbegrenztem) Speicher.



Auf den Keller (Stack, Stapel) kann ein Element an oberster Stelle zugefügt werden (*push*) oder es kann das oberste Element (*pop*) entfernt werden.

Kelleroperationen:



Beispiel (Informell)

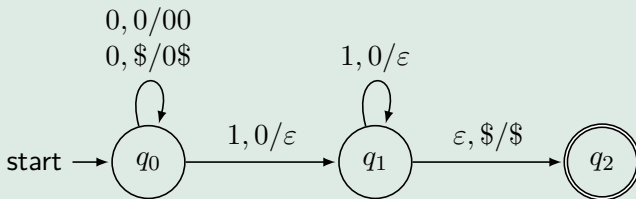
Ein Kellerautomat für die kontextfreie Sprache $\{0^n 1^n \mid n > 0\}$:

- Solange keine Eins gelesen wird, lege die gelesenen Nullen auf dem Keller ab. Sobald Einsen gelesen werden, entferne für jede gelesene Eins eine Null vom Keller.
- Akzeptiere das Eingabewort, wenn die Berechnung im akzeptierenden Zustand endet. Der akzeptierende Zustand wird erreicht, wenn der Keller leer ist und das ganze Wort gelesen wurde.
- Andernfalls verwirfe die Eingabe.

Anmerkung: engl. pushdown automata (PDA)

Beispiel (Grafisch)

Der Automat als Diagramm.



Anmerkung: Das Zeichen \$ zeigt an, dass der Stack leer ist.

Definition (deterministischer Kellerautomat)

Ein **deterministischer Kellerautomat (KA)** ist ein 7-Tupel $(Q, \Sigma, \Gamma, \delta, q_0, \$, F)$, wobei

- Q ist eine endliche Menge von Zuständen.
- Σ ist das Alphabet der Eingabe.
- Γ ist das Alphabet des Kellers.
- $\delta : Q \times (\Sigma \cup \varepsilon) \times \Gamma \rightarrow Q \times \Gamma^*$ ist eine (partielle) Übergangsfunktion.
- $q_0 \in Q$ ist der Startzustand.
- $\$ \in \Gamma$ ist ein ausgezeichnetes Symbol vom Alphabet des Kellers.
- $F \subseteq Q$ ist die Menge der akzeptierenden Zustände.

Anmerkung: Anfangs enthält der Keller eine Instanz des Symbols \$.

Definition (Fortsetzung)

Für die Übergangsfunktion gilt zusätzlich folgende Einschränkung:

Für jeden Zustand q und alle Symbole x, b gilt, wenn $\delta(q, b, x)$ definiert ist, dann ist $\delta(q, \varepsilon, x)$ undefiniert.

Anmerkung: Diese Bedingung ist nötig, um sicherzustellen, dass jeder KA tatsächlich deterministisch ist.

Beispiel (Formale Beschreibung)

Ein KA für die kontextfreie Sprache $\{0^n 1^n \mid n > 0\}$:

Formal lässt sich der Automat als $(Q, \Sigma, \Gamma, \delta, q_0, \$, F)$ darstellen, wobei

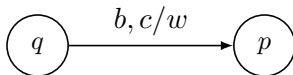
- $Q = \{q_0, q_1, q_2\}$,
- $\Sigma = \{0, 1\}, \Gamma = \{0, \$\}$,
- $F = \{q_2\}$,
- und δ wie folgt gegeben ist:

$$\begin{aligned} \delta(q_0, 0, 0) &= (q_0, 00) & \delta(q_0, 0, \$) &= (q_0, 0\$) & \delta(q_0, 1, 0) &= (q_1, \varepsilon) \\ \delta(q_1, 1, 0) &= (q_1, \varepsilon) & \delta(q_1, \varepsilon, \$) &= (q_2, \$). \end{aligned}$$

Ein Berechnungsschritt $\delta(q, b, c) = (p, w)$ wird wie folgt interpretiert:

- Der Automat befindet sich im Zustand q .
- Der Automat liest das Symbol b von der Eingabe (falls $b = \varepsilon$, wird nichts gelesen).
- Der Automat entfernt das oberste Kellersymbol c .
- Der Automat schreibt das Wort w auf den Stack (von hinten nach vorne).
- Der Automat wechselt in den Zustand p .

Ein Übergang $\delta(q, b, c) = (p, w)$ wird graphisch als



dargestellt. Analog zu den endlichen Zustandsautomaten gelten folgende Konventionen:

- Akzeptierende Zustände werden mit einer doppelten Konturlinie gekennzeichnet.
- Der Anfangszustand wird durch einen eingehenden Pfeil gekennzeichnet.

Definition (nichtdeterministischer Kellerautomat)

Ein **nichtdeterministischer Kellerautomat (NKA)** ist ein 7-Tupel $(Q, \Sigma, \Gamma, \delta, q_0, \$, F)$, der sich vom KA nur in der Definition der Übergangsfunktion unterscheidet:

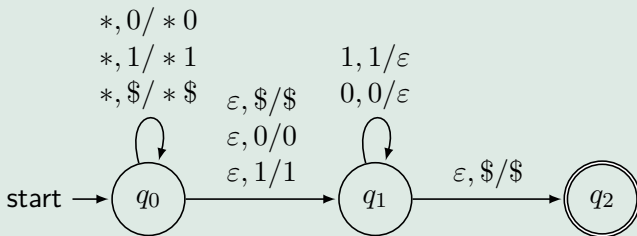
$$\delta : Q \times (\Sigma \cup \varepsilon) \times \Gamma \rightarrow \mathcal{P}(Q \times \Gamma^*).$$

Anmerkung:

- Die Zusatzbedingung an die Übergangsfunktion fällt beim NKA weg.
- Analog zum NEA bildet die Übergangsfunktion des NKA in die Potenzmenge ab.

Beispiel

Kellerautomat für die Sprache $\{ ww^R \mid w \in \{0, 1\}^* \}$:



Anmerkung:

- w^R ist das Wort w rückwärts geschrieben.
- Der $*$ steht im Diagramm für ein beliebiges Zeichen aus Σ (d. h. in diesem Beispiel 0 oder 1) und dient der Vereinfachung.

Definition (Konfiguration)

Sei $M = (Q, \Sigma, \Gamma, \delta, q_0, \$, F)$ ein NKA.

Eine **Konfiguration von** M ist ein Element (q, w, γ) aus $Q \times \Sigma^* \times \Gamma^*$, wobei

- q für den Zustand steht,
- w die verbleibende Eingabe repräsentiert,
- γ für den Inhalt des Kellers steht.
(Dabei ist das Symbol ganz links das oberste Symbol)

Mit $(q_0, w, \$)$ bezeichnen wir die **Startkonfiguration** für die Eingabe w und mit (q, ε, γ) eine **Endkonfiguration**.

Definition (Berechnungsschritt)

Sei $M = (Q, \Sigma, \Gamma, \delta, q_0, \$, F)$ ein NKA. Seien $w \in \Sigma^*$ und $\gamma \in \Gamma^*$.

Ein **Berechnungsschritt** \vdash von M ist die Anwendung der Übergangsfunktion auf die aktuelle Konfiguration und ist definiert durch

$$(q, aw, b\gamma) \vdash (p, w, u\gamma)$$

genau dann, wenn $(p, u) \in \delta(q, a, b)$.

Für zwei Konfigurationen K und K' schreiben wir $K \vdash^* K'$, falls es weitere Konfigurationen K_1, \dots, K_n gibt mit

$$K \vdash K_1 \vdash \dots \vdash K_n \vdash K'.$$

Definition (Berechnung)

Sei $M = (Q, \Sigma, \Gamma, \delta, q_0, \$, F)$ ein NKA. Seien $w \in \Sigma^*$ und $\gamma \in \Gamma^*$.

Eine **Berechnung von M auf w** ist eine Folge von Berechnungsschritten, die in der Startkonfiguration beginnt und in einer Endkonfiguration $(q_f, \varepsilon, \gamma)$ endet, von der aus kein weiterer Berechnungsschritt mehr möglich ist.

Die Berechnung ist **akzeptierend**, wenn für die Endkonfiguration $(q_f, \varepsilon, \gamma)$ gilt, dass $q_f \in F$.

Beispiel

Berechnung für $w' = 0011$ u. $w'' = 011$ für $L = \{0^n 1^n \mid n > 0\}$

- Für w' : $(q_0, 0011, \$) \vdash (q_0, 011, 0\$) \vdash (q_0, 11, 00\$) \vdash (q_1, 1, 0\$)$
 $\vdash (q_1, \epsilon, \$) \vdash (q_2, \epsilon, \$)$

\Rightarrow Die Berechnung ist akzeptierend.

- Für w'' : $(q_0, 011, \$) \vdash (q_0, 11, 0\$) \vdash (q_1, 1, \$)$

\Rightarrow Die Berechnung ist nicht akzeptierend.

Definition (Sprache $L(M)$)

Sei $M = (Q, \Sigma, \Gamma, \delta, q_0, \$, F)$ ein NKA.

Die **Sprache** $L(M)$ des Kellerautomaten M ist definiert durch

$$L(M) = \{ w \in \Sigma^* \mid (q_0, w, \$) \vdash^* (q, \varepsilon, \gamma) \text{ für ein } q \in F \text{ und ein } \gamma \in \Gamma^* \}.$$

Elemente von $L(M)$ werden (von M) **akzeptierte** Wörter genannt.

Anmerkung: Ein Wort w wird genau dann von M akzeptiert, wenn es eine akzeptierende Berechnung von M auf w gibt ("*Der KA akzeptiert durch Endzustand*"¹).

¹In der Lit. wird auch die Akzeptanz durch einen leeren Keller verwendet. Beide Definitionen sind für den NKA äquivalent (nicht aber für den DKA).

Theorem (kontextfreie Sprache)

Eine Sprache ist kontextfrei, genau dann, wenn es einen nichtdeterministischen Kellerautomaten gibt, der die Sprache erkennt.

Beweis.

Siehe Hopcroft et al. S. 248 ff.



Anmerkungen:

- Es gibt kontextfreie Sprachen, die nicht von einem KA erkannt werden. Insbesondere wird nicht jede Sprache, die von einem NKA erkannt wird, auch von einem KA erkannt.
- Kontextfreie Sprachen, die von einem KA erkannt werden sind dann auch eindeutig. Sie spielen bei der Syntax-Analyse (Parser) eine grosse Rolle.

Das folgende Beispiel zeigt, dass es auch Sprachen gibt, die von keinem Kellerautomaten erkannt werden können.

Beispiel

Die Sprache $L = \{ 0^n 1^n 2^n \mid n > 0 \}$ ist nicht kontextfrei.

- Ein Kellerautomat, der ein Wort aus L akzeptieren würde, müsste sich die Anzahl der eingelesenen 0 und 1 merken.
- Mit dem Keller kann aber nur *einmal* eine Anzahl verglichen werden, danach sind die Symbole nicht mehr auf dem Keller.
- Mit Zuständen kann, wie beim EA, nicht beliebig gezählt werden.