

Vorlesung Baumautomaten (Mitschrift)

Benedikt Elßmann (3720358)
be57xocu@studserv.uni-leipzig.de

Universität Leipzig

11. April 2019

Inhaltsverzeichnis

0	Einleitung	2
1	Bäume und Baumautomaten	2
1.1	Definition Rangalphabet	3
1.2	Definition Term, Tree	3
1.3	Definition Höhe	4
1.4	Definition Position	4
1.5	Definition der Label an den Positionen	4
1.6	Definition Sub-Baum	4
1.7	Definition Baumautomat	5
1.8	Definition Lauf/Run	5
1.9	Lemma	6

0 Einleitung

Automaten lesen Wörter $w = a_1 \dots a_n$ und geben "accept" aus oder nicht. Dafür gibt es Erweiterungen, wie etwa:

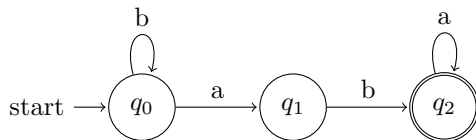
- gewichtete Automaten, das heißt der Output ist ein Semiringelement
- Automaten mit Gedächtnis (Stack)
- Automaten über anderen Strukturen
 - ω -Wörter $w = a_1 \dots a_n$
 - Graphen
 - Bäume
 - Kombinationen dieser

Typische Fragestellungen:

- Ausdrucksstärke
- Darstellung als rationale Ausdrücke (Kleene)
- Darstellung als Grammatik
- Darstellung als Logik

1 Bäume und Baumautomaten

Wir betrachten über $A = \{a, b\}$ den Automaten \mathcal{A} :



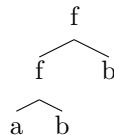
mit $L(\mathcal{A}) = b^*aba^*$.

Betrachtung des Wortes $w = baba \in L(\mathcal{A})$:

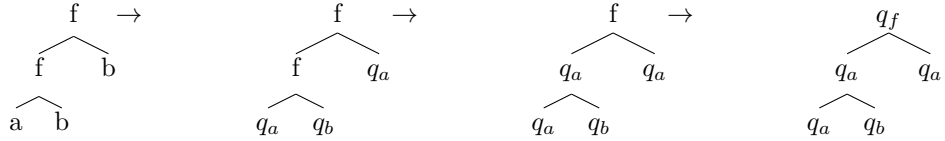
Der eindeutige erfolgreiche Lauf für w lässt sich darstellen als:

$q_0baba \rightarrow bq_0aba \rightarrow baq_1ba \rightarrow babq_2a \rightarrow babaq_2 \in F$ (Finalzustand)

Baumautomaten funktionieren analog. Unser erstes Beispiel wird



Akzeptiert mit dem Lauf:



mit $q_f \in F$

1.1 Definition Rangalphabet

Ein paar (Σ, rk) , wobei Σ eine endliche Menge von Symbolen und $rk : \Sigma \rightarrow \mathbb{N}$ eine Abbildung ist, heißt Rangalphabet.

Für $f \in \Sigma$ heißt $rk(f)$ der Rang (oder die Stelligkeit) von f .

Intuitiv: $rk(f)$ ist die Anzahl der Kinder von f in einem Baum.

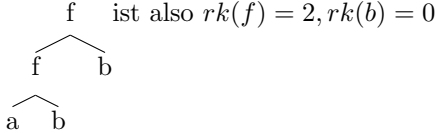
Insbesondere ist die Anzahl der Kinder für jedes Symbol fest.

Gilt $rk(f) = n$, schreiben wir auch $f^{(n)}$ statt f . wir schreiben:

- 0-stellige Symbole (Konstanten) a, b, \dots
- unär, binär, $\dots f, g, \dots$

Wir setzen $\Sigma^{(n)} = \{f \in \Sigma \mid rk(f) = n\}$

In



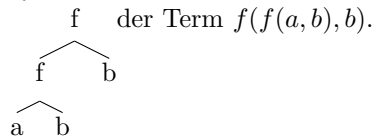
1.2 Definition Term, Tree

Sei (Σ, rk) ein Rangalphabet. Die Menge T_Σ der Bäume über Σ ist induktiv definiert durch:

- $\Sigma^0 \subseteq T_\Sigma$
- $f^{(n)} \in \Sigma$. $t_1, \dots, t_n \in T_\Sigma$, dann ist $f(t_1, \dots, t_n) \in T_\Sigma$

Intuitiv sind t_1, \dots, t_n die Kinder von f .

Z.B. ist



1.3 Definition Höhe

Sei (Σ, rk) ein Rangalphabet. Die Höhe ht ist gegeben durch:

- für $a^{(0)} \in \Sigma : ht(a) = 1.$
- für $f(t_1, \dots, t_n) \in T_\Sigma : ht(f) = 1 + \max\{ht(t_i) | i \in \{1, \dots, n\}\}$

Ziel: Zugriff auf einen Knoten innerhalb eines Baumes und deren Label.
Dafür ordnen wir den Knoten Positionen zu. Das geht induktiv wie folgt:

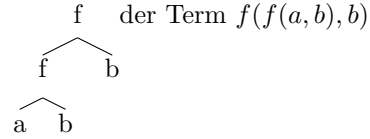
1.4 Definition Position

Sei (Σ, rk) ein Rangalphabet. Die Positionenmenge ist definiert durch:

- für $a^{(0)} \in T_\Sigma$ ist $Pos(a) = \{\varepsilon\}$
- für $f(t_1, \dots, t_n) \in T_\Sigma$ ist $Pos(f(t_1, \dots, t_n)) = \{\varepsilon\}1 \cdot Pos(t_1) \cup \dots \cup n \cdot Pos(t_n)$

Beispiel:

Betrachtung von $f(f(a, b), b)$ bzw.



$$Pos(f) = \{\varepsilon, 1, 2, 1.1, 1.2\}$$

1.5 Definition der Label an den Positionen

Für einen Term der Form $t = f(t_1, \dots, t_n)$ ist das Symbol $t(p)$ in t an p-ter Position induktiv definiert durch:

- $t(\varepsilon) = f$
- $t(ip) = t_i(p), i \in \{1, \dots, n\}$

Beispiel: Betrachtung von $f(f(a, b), b)$

Dann ist

$$t(\varepsilon) = f$$

$$t(1) = t(1 \cdot \varepsilon) = t_1(\varepsilon) = f$$

$$t(2) = t(2 \cdot \varepsilon) = t_2(\varepsilon) = b$$

$$t(1.1) = t_1(1) = a$$

$$t(1.2) = t_1(2) = b$$

1.6 Definition Sub-Baum

Für T_Σ ist ein Sub-Baum $t|_p$ an p-ter Position wie folgt definiert:

- $Pos(t|_p) = \{i|pi \in Pos(t)\}$

- $\forall q \in Pos(t|_p)$ ist $t|_p(q) = t(pq)$

Wir schreiben $t[u]_p$ für den Baum, der entsteht, wenn man in t den sub-Baum $t|_p$ durch u ersetzt.

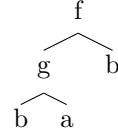
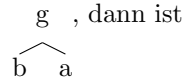
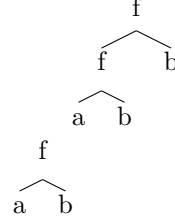
Beispiel: $f(f(a, b), b)$ bzw.

$$t|_1 = f(a, b)$$

$$t|_2 = t(1.2) = b$$

$$u = g(b, a)$$

$$t[u]_1 = f(g(b, a), b)$$



1.7 Definition Baumautomat

Ein Baumautomat \mathcal{A} ist ein 4-Tupel (Q, Σ, F, Δ) , wobei:

$Q \dots$ endliche Menge an Zuständen

$\Sigma \dots$ Rangalphabet, wobei $\Sigma \cup Q \neq \emptyset$

$F \dots \subseteq Q$ Finalzustände

$\Delta \dots$ Menge von Regeln

$$r : f(q_1 \dots q_n) \rightarrow q$$

für $q, q_1, \dots, q_n \in Q$, für $a^{(0)} \in T_\Sigma : a \rightarrow q$

Beispiel:

$$\mathcal{A} = \{\{q_a, q_b, q_f\}, \{a^{(0)}, b^{(0)}, f^{(2)}\}, \{q_f\}, \Delta\}$$

$$\text{mit } \Delta = \{a \rightarrow q_a, b \rightarrow q_b, f(q_a, q_b) \rightarrow q_a, f(q_a, q_b), f(q_a, q_b) \rightarrow q_f\}$$

1.8 Definition Lauf/Run

Sei $\mathcal{A} = (Q, \Sigma, F, \Delta)$ ein Baumautomat und $t \in T_\Sigma$. Ein Lauf r für t von \mathcal{A} ist ein Term mit

- $Pos(r) = Pos(t)$
- Ist $t(p) = a$ ein Blatt, dann ist $r(p) = q_a$, nur wenn $(a \rightarrow q_a) \in \Delta$
- Ist $t(p) = f^{(m)}$, dann ist $r(p) = q$, wenn $(f(q_1, \dots, q_n) \rightarrow q) \in \Delta$ und $r(p_i) = q_i, i \in \{1, \dots, n\}$

Ein Lauf ist erfolgreich, wenn $r(\varepsilon) \in F$. Der Automat \mathcal{A} akzeptiert t , falls es einen erfolgreichen Lauf für t von \mathcal{A} gibt.

Wir bezeichnen mit $L(\mathcal{A}) = \{t \in T_\Sigma \mid \mathcal{A} \text{ akzeptiert } t\}$ die von \mathcal{A} erkannte Baumsprache. Eine Sprache $L \subseteq T_\Sigma$ heißt erkennbar, falls ein Baumautomat \mathcal{A} existiert mit $L = L(\mathcal{A})$.

Um einzelne Schritte von Baumautomaten zu formalisieren, betrachten wir die *move relation* $\rightarrow_{\mathcal{A}}$, definiert wie folgt:

Gegeben sei $\mathcal{A} = (Q, \Sigma, F, \Delta)$, dann ist $t \rightarrow_{\mathcal{A}} t'$ mit $t, t' \in T_{\Sigma \cup Q}$, falls

- $t(p) = f^{(n)}$
- $t(pi) = q_i$ für $i \in \{1, \dots, n\}$ und p_i sind Blätter
- $(f(q_1, \dots, q_n) \rightarrow q) \in \Delta$
- und $t' = t[q]_p$

Mit $\rightarrow_{\mathcal{A}}^*$ bezeichnen wir die transitive Hülle von $\rightarrow_{\mathcal{A}}$.

1.9 Lemma

Sei $\mathcal{A} = (Q, \Sigma, F, \Delta)$ ein Baumautomat. Dann ist $L(\mathcal{A}) = \{t \in T_\Sigma \mid t \rightarrow_{\mathcal{A}}^* q \text{ mit } q \in F\} (= Z)$

Beweis: „ $L(\mathcal{A}) \subseteq Z$ “:

Wir zeigen: Es existiert ein Run r für t von \mathcal{A} mit $r(\varepsilon) = q$, dann ist $t \rightarrow_{\mathcal{A}}^* q$

Inuktionsannahme:

$t = a^{(0)} \in T_\Sigma$. Dann gilt $a \in L(\mathcal{A})$, falls ein Lauf r existiert mit $r(a) = q_a$ und $(a \rightarrow q_a) \in \Delta$. Dann folgt $a \rightarrow_{\mathcal{A}}^* q_a$.

Sei nun $t = f(t_1, \dots, t_n)$

Induktionsvoraussetzung:

Falls für t_1, \dots, t_n Läufe r_i existieren mit $r_i(\varepsilon) = q_i$, dann gilt auch $t_i \rightarrow_{\mathcal{A}}^* q_i$ mit $i \in \{1, \dots, n\}$

Induktionsschritt:

zu zeigen: Es existiert ein Lauf r für t mit $r(\varepsilon) = q$, dann $t \rightarrow_{\mathcal{A}}^* q$.

Sei also r ein Lauf mit $r(\varepsilon) = q$. Dann ist $r(i) = q_i$, $i \in \{1, \dots, n\}$, mit $(f(q_1, \dots, q_n) \rightarrow q) \in \Delta$.

Laut Induktionsvoraussetzung gilt nun, $t_i \rightarrow_{\mathcal{A}}^* q_i$, $i \in \{1, \dots, n\}$.

Damit $t = f(t_1, \dots, t_n) \rightarrow_{\mathcal{A}}^* f(q_1, t_2, \dots, t_n) \rightarrow_{\mathcal{A}}^* \dots \rightarrow_{\mathcal{A}}^* f(q_1, \dots, q_n)$

Des weiteren haben wir die regel $f(q_1, \dots, q_n) \rightarrow q$, das heißt $f(q_1, \dots, q_n) \rightarrow_{\mathcal{A}}^* q$.

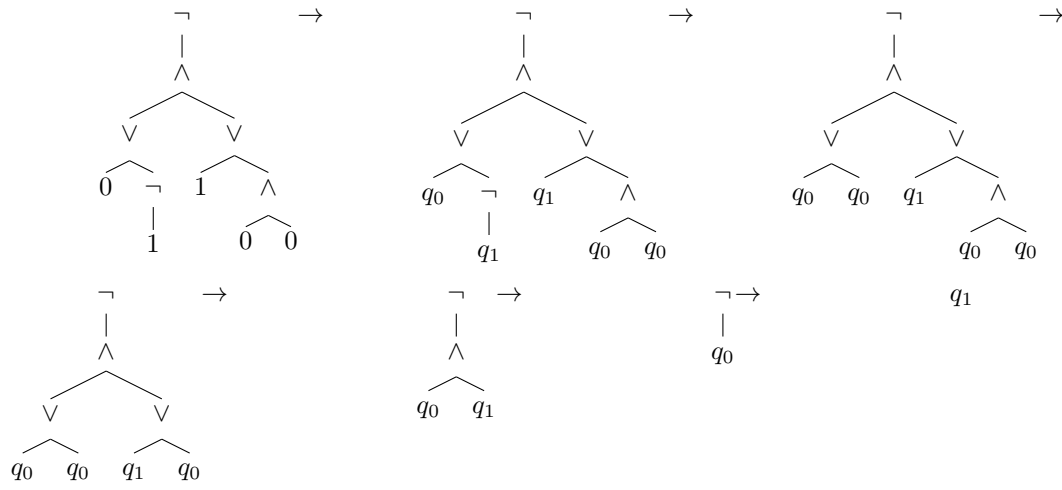
Insgesamt also $t \rightarrow_{\mathcal{A}}^* q$

Beweis: „ $L(Z \subseteq \mathcal{A})$ “: analog

Einige Beispiele für Baumentautomaten:

- Sei $B = (\{q_0, q_1\}, \{0^{(0)}, 1^{(0)}, \neg^{(1)}, \wedge^{(2)}, \vee^{(2)}\} \{q_1\}, \Delta)$ mit
 $\Delta = \{0 \rightarrow q_0, 1 \rightarrow q_1,$
 $\neg(q_0) \rightarrow q_1, \neg(q_1) \rightarrow q_0,$
 $\wedge(q_0, q_0) \rightarrow q_0, \wedge(q_0, q_1) \rightarrow q_0, \wedge(q_1, q_0) \rightarrow q_0, \wedge(q_1, q_1) \rightarrow q_1$
 $\vee(q_0, q_0) \rightarrow q_0, \vee(q_0, q_1) \rightarrow q_1, \vee(q_1, q_0) \rightarrow q_1, \vee(q_1, q_1) \rightarrow q_1\}$

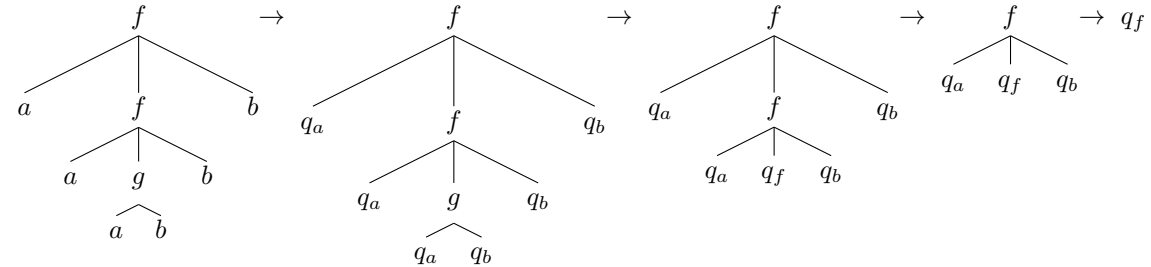
Beispiellauf:



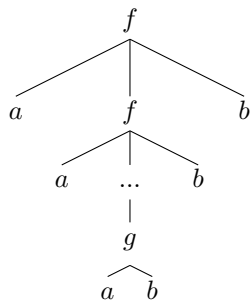
- $(a^n b^n \text{light})$

Betrachten $\mathcal{A} = (\{q_a, q_b, q_f\}, \{a^{(0)}, b^{(0)}, f^{(3)}, g^{(2)}\}, \{q_f\}, \Delta)$
mit $\Delta =$
 $\{a \rightarrow q_a, b \rightarrow q_b, g(q_a, q_b) \rightarrow q_f, f(q_a, q_f, q_b) \rightarrow q_f\}$

Beispiellauf:



\mathcal{A} akzeptiert also alle Bäume der Form:



3. Simulation eines Wortautomaten: (siehe Übung)