Endliche Akzeptoren (14)

Definition: Der Akzeptor a heißt **vollständig** $:\Longleftrightarrow$ es gelten

(1) $\mu(s) \in X \cup \{1\}$ für alle Kanten s. Aus $\mu(s) = \mu(s') = 1$ folgt $s = s'$ und $Q(s) \in S, Z(s) \in T$.

(2) Zu jedem Knoten P und jedem $x \in X$ gibt es eine Kante s mit $Q(s) = P, \mu(s) = x$.

Lemma: Zu jedem Akzeptor a gibt es einen vollständigen Akzeptor a' mit $L_a = L_{a'}$.

Beweis: Falls a nicht vollständig ist, führe einen neuen Knoten P_∞ ein und von jedem Knoten P aus, von dem keine Kante s mit $\mu(s) = x$ ausgeht, eine Kante

$$P \xrightarrow{x} P_\infty$$

Endliche Akzeptoren (15)

Weiter auch Kanten

$$P_\infty \xrightarrow{x} P_\infty$$

für $x \in X$ (d.h. P_∞ kann man nicht verlassen). Die Iteration dieses Schrittes führt zu einem vollständigen Automaten a' mit $L_{a'} = L_a$.

Definition: a heißt deterministischer Akzeptor $:\Longleftrightarrow$ zu jedem Knoten P und $x \in X$ gibt es genau eine Kante s mit $Q(s) = P$ und $\mu(s) = x$.

Lemma: Zu jedem Akzeptor a gibt es einen vollständigen deterministischen Akzeptor $\tilde{a}$ mit $L_a = L_{\tilde{a}}$.

Beweis: Wir dürfen annehmen, daß a vollständig ist und daß $\mu(s) \in X$ für $s \in E(G)$ gilt.

Endliche Akzeptoren (16)

Konstruktion von $\tilde{a}$.

Wir setzen

$$\tilde{K} := 2^K.$$

Sind $Z, Z' \in \tilde{K}$, so gibt es genau dann eine Kante $\tilde{s}$ mit $\tilde{\mu}(\tilde{s}) = x \in X$, wenn es zu jedem $z' \in Z'$ eine Kante $s \in K$ gibt mit $Q(s) \in Z$, $Z(s) = z'$ und $\mu(s) = x$. Da wir voraussetzen, daß a vollständig ist, gibt es dann auch zu jedem $z \in Z$ eine Kante s mit $Q(s) = z$ und $Z(s) \in Z'$ und $\mu(s) = x$.

$$Z = \left\{ \begin{array}{c} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_k \end{array} \right\} \xrightarrow[\tilde{s}]{x} \left\{ \begin{array}{c} z'_1 \\ \vdots \\ z'_j \\ \vdots \\ z'_l \end{array} \right\} = Z'$$

Endliche Akzeptoren (17)

Es gibt genau dann eine Kante $\tilde{s}$

$$Q(\tilde{s}) = S \in \tilde{K},\ Z(\tilde{s}) = T \in \tilde{K},\ \tilde{\mu}(\tilde{s}) = 1,$$

wenn es eine Kante s in E gibt mit $Q(s) \in S$, $Z(s) \in T$ und $\mu(s) = 1$.

Damit haben wir den Graphen $\tilde{G}$ und die Markierung $\tilde{\mu}$ definiert. Wir setzen weiter

$$\tilde{S} := S \text{ und } \tilde{T} = \{Z \in \tilde{K} \mid Z \cap T \neq \emptyset\}$$

und

$$\tilde{a} := (\mathcal{W}(\tilde{G}), \tilde{S}, \tilde{T}, \tilde{\mu}, X)$$

und verwenden die Bezeichnung $2^a := \tilde{a}$.

Endliche Akzeptoren (18)

Wir beweisen $L_{\tilde{a}} = L_a$.

(1) Zu jedem $u \in L_a$ gibt es ein $w \in \mathcal{W}(G)(S,T)$ mit $\mu(w) = u$. Es gibt also $s \in S$ und $t \in T$ mit

$$s \xrightarrow[s_1]{x_1} z_1 \xrightarrow[s_2]{x_2} z_2 \xrightarrow[s_3]{x_3} \ \ldots \ \xrightarrow[s_n]{x_n} t,$$

worin $u = x_1 \cdot \ldots \cdot x_n$ und $\mu(s_i) = x_i$ ist. Hieraus folgt unmittelbar, daß es einen Weg $\tilde{w} = (\tilde{s}_1, \ldots, \tilde{s}_n) \in \mathcal{W}(\tilde{G})(\tilde{S}, \tilde{T})$ mit $\tilde{\mu}(\tilde{w}) = u$ gibt. Also ist $L_a \subset L_{\tilde{a}}$.

Endliche Akzeptoren (19)

(2) Zu $u \in L_{\tilde{a}}$, $|u| = n$, gibt es einen Weg $\tilde{w} = (\tilde{s}_1, \ldots, \tilde{s}_n) \in \mathcal{W}(\tilde{G})(\tilde{S}, \tilde{T})$ mit $\tilde{\mu}(\tilde{w}) = u$. Da a als vollständig vorausgesetzt wird, gibt es auch zu jedem $P \in S$ einen Weg $w = (s_1, \ldots, s_n)$ in $\mathcal{W}(G)$ mit $Q(w) = P$, und $\mu(w) = u$. Aufgrund der Definition von $\tilde{a}$ folgt $Z(w) \in Z(\tilde{w}) \in \tilde{T}$. Es braucht aber **nicht** $z(w) \in T$ zu gelten! Wir finden einen geeigneten Weg w mit $z(w)$, indem wir $\tilde{z} \in T \cap Z(\tilde{w})$ auswählen. Nun gibt es aufgrund der Definition von $\tilde{G}$ eine Kante s_n mit $\mu(s_n) = x$ und $Z(s_n) = \tilde{z}$. Ebenso findet man s_{n-1} als Kante mit $\mu(s_{n-1}) = x_{n-1}$ und $Z(s_{n-1}) = Q(s_n) \in Z(\tilde{s}_{n-1})$ usw. Der so zu $\tilde{w}$ konstruierte Weg w liegt in $\mathcal{W}(G)(S, T)$ und erfüllt die Forderung $\mu(w) = u$. Also ist $u \in L_a$ und also $L_{\tilde{a}} \subset L_a$.

Aus (1) und (2) folgt $L_a = L_{\tilde{a}}$.

Endliche Akzeptoren (20)

Bemerkung: Die Konstruktion hat $\#S = 1$ nicht verwendet und auch nicht, daß es keine Kante gibt, die in S einmündet. Unter Verwendung dieser Eigenschaft gilt für jeden Weg $w \in \mathcal{W}(\tilde{G})(\tilde{S}, Z)$ mit $|w| \geq 1$, daß $Z(w) \cap \tilde{S} = \emptyset$ ist. Das erlaubt es uns, die im Falle $1 \in L_{\tilde{a}}$ vorhandene $1-$Kante zu eliminieren, indem wir $\tilde{T}$ durch $\tilde{T} \cup \tilde{S}$ ersetzen.

Übung: Kontrolliere, ob sich dieser Beweis ohne die Voraussetzung der Vollständigkeit von a ebenso durchführen läßt.

Endliche Akzeptoren (21)

Lemma: Sei a vollständig und deterministisch und sei $\tilde{a}$ ein Akzeptor, der sich von a nur dadurch unterscheidet, daß $\tilde{T} = K - T$ ist. Es gilt dann $L_{\tilde{a}} = X^* - L_a$.

Beweis: Ist $u \in L_a$, dann gibt es in a einen Weg $w \in \mathcal{W}(G)(S,T)$ mit $\mu(w) = u$ mit $Z(w) \in T$. Da a deterministisch ist, ist w der einzige solche Weg. Also gibt es keinen Weg $w' \in \mathcal{W}(G)(S, K-T)$ mit $\mu(w') = u$. Also ist $u \notin L_{\tilde{a}}$. Da die Definitionen von a und $\tilde{a}$ bezüglich der Komplementbildung symmetrisch sind, gilt auch umgekehrt $u \in L_{\tilde{a}} \Rightarrow u \notin L_a$. Also ist $L_{\tilde{a}} = X^* - L_a$.

Fassen wir die gesamte Konstruktion ins Auge, dann gilt also: Zu jedem endlichen Akzeptor a gibt es einen vollständigen deterministischen Akzeptor $\tilde{a}$ mit $L_{\tilde{a}} = X^* - L_a$, womit unser Satz auf Blatt 26 bewiesen ist.

Endliche Akzeptoren (22)

Beobachtung: Setzen wir

$$|a| = \#K,$$

dann gilt für unsere Konstruktion

$$|\tilde{a}| = 2^{|a|}.$$

Es mag aber nun sein, daß es zu a einen deterministischen Akzeptor $\tilde{a}$ gibt, mit

$$|a| = |\tilde{a}|.$$

Das ist der Fall, wenn bereits a vollständig und deterministisch ist. In diesem Fall erzeugt unsere Konstruktion sehr viele Zustände, die von keinem von $\tilde{S}$ ausgehenden Weg erreicht werden. Diese Zustände sind also überflüssig. Hierdurch stellt sich die

Frage: Gibt es eine Folge $a_1, a_2, a_3, \ldots$ von Akzeptoren über dem Alphabet X, so daß für jede Folge $\tilde{a}_1, \tilde{a}_2, \tilde{a}_3, \ldots$ von vollständigen deterministischen Akzeptoren mit $L_{a_i} = L_{\tilde{a}_i}$ für $i = 1, 2, 3, \ldots$

$$|\tilde{a}_i| \approx 2^{|a_i|}$$

gilt.

Endliche Akzeptoren (23)

Wir betrachten hierzu ein Beispiel von M. Fischer. (Siehe: G. Hotz: Einführung in die Informatik, Teubner-Verlag)

Sei

$$L_n = \{u \cdot v \in \{a,b\}^* \mid |u| = |v| = n, u \neq v\}.$$

Wir wollen die Akzeptoren aus einfacheren Akzeptoren aufbauen. Hierzu definieren wir

$$\begin{aligned} L_n^i = \{&u_1 a u_2 v_1 b v_2,\ u_1 b u_2 v_1 a v_2 \mid \\ &i = |u_1| = |v_1|, |u_2| = |v_2| = n - i - 1\} \end{aligned}$$

Es gilt offensichtlich

$$L_n = L_n^0 \cup L_n^1 \cup \cdots \cup L_n^{n-1}$$

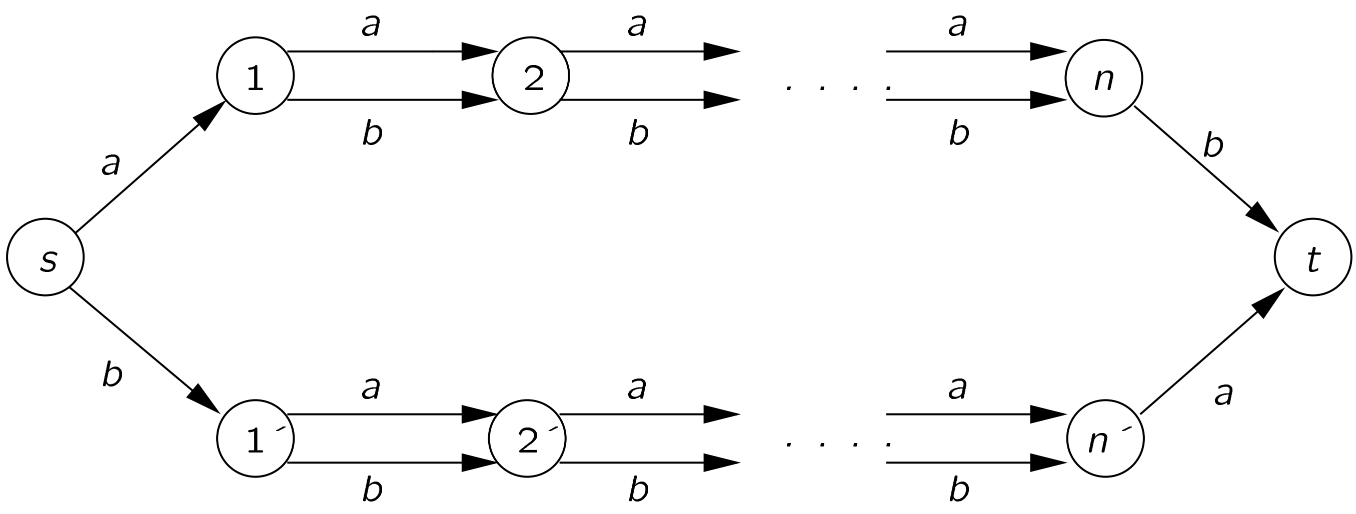
Endliche Akzeptoren (24)

Weiter gilt für $i = 0, \ldots, n-1$

$$\begin{aligned} L_n^i &= \{a,b\}^i \cdot a \cdot \{a,b\}^{n-1} \cdot b \cdot \{a,b\}^{n-i-1} \\ &\cup \ \{a,b\}^i \cdot b \cdot \{a,b\}^{n-1} \cdot b \cdot \{a,b\}^{n-i-1} \\ \\ &= \{a,b\}^i \cdot \Big(a \cdot \{a,b\}^{n-1} \cdot b \ \cup \\ &\quad b \cdot \{a,b\}^{n-1} \cdot a \Big) \cdot \{a,b\}^{n-i-1} \end{aligned}$$

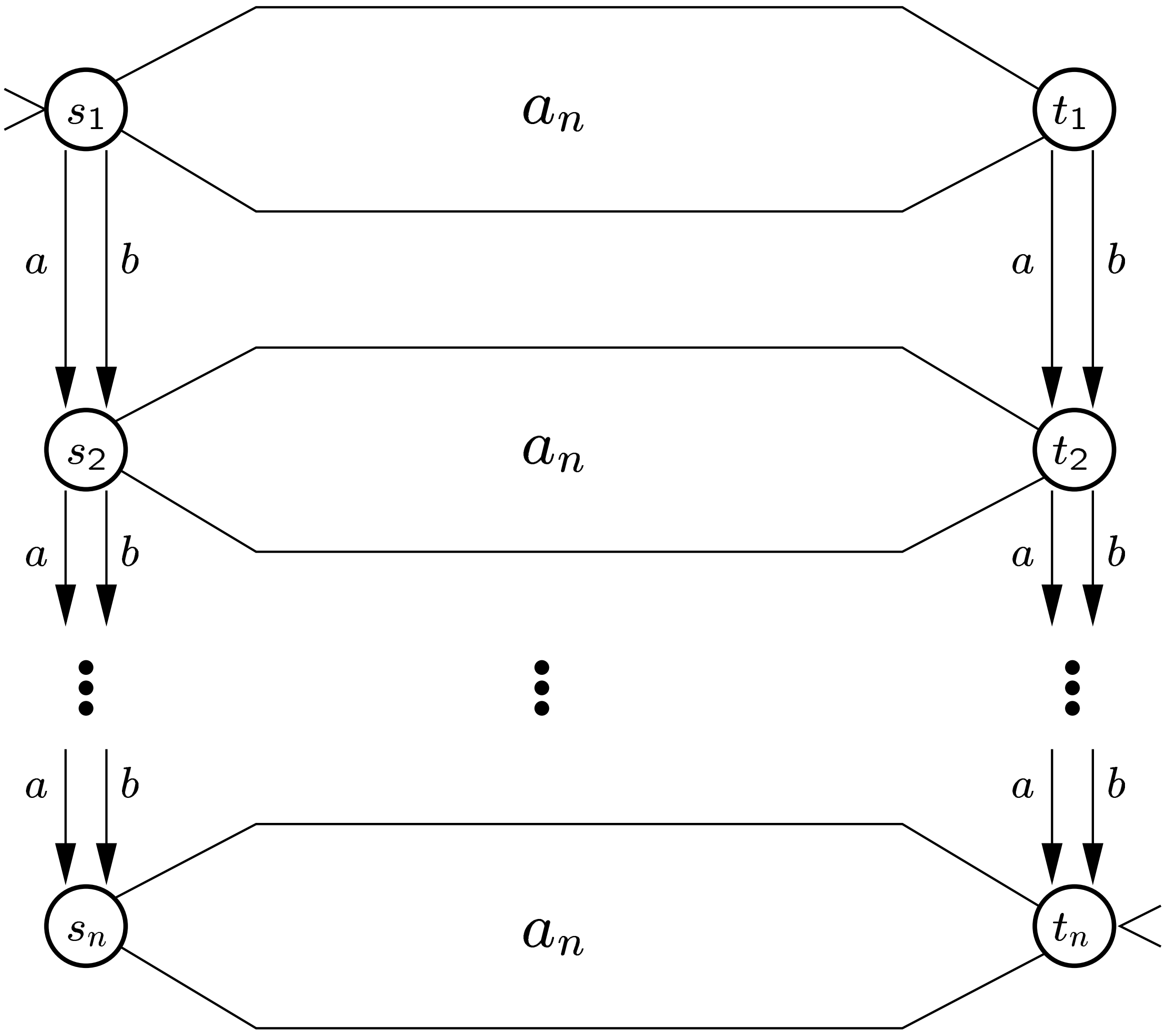
Weiter gilt natürlich $\{a,b\}^{i+1} = \{a,b\}^i \cdot \{a,b\}$.

Wir geben zunächst einen Akzeptor a_n für $a \cdot \{a,b\}^{n-1} \cdot b \cup b \cdot \{a,b\}^{n-1} \cdot a$ an:



Endliche Akzeptoren (25)

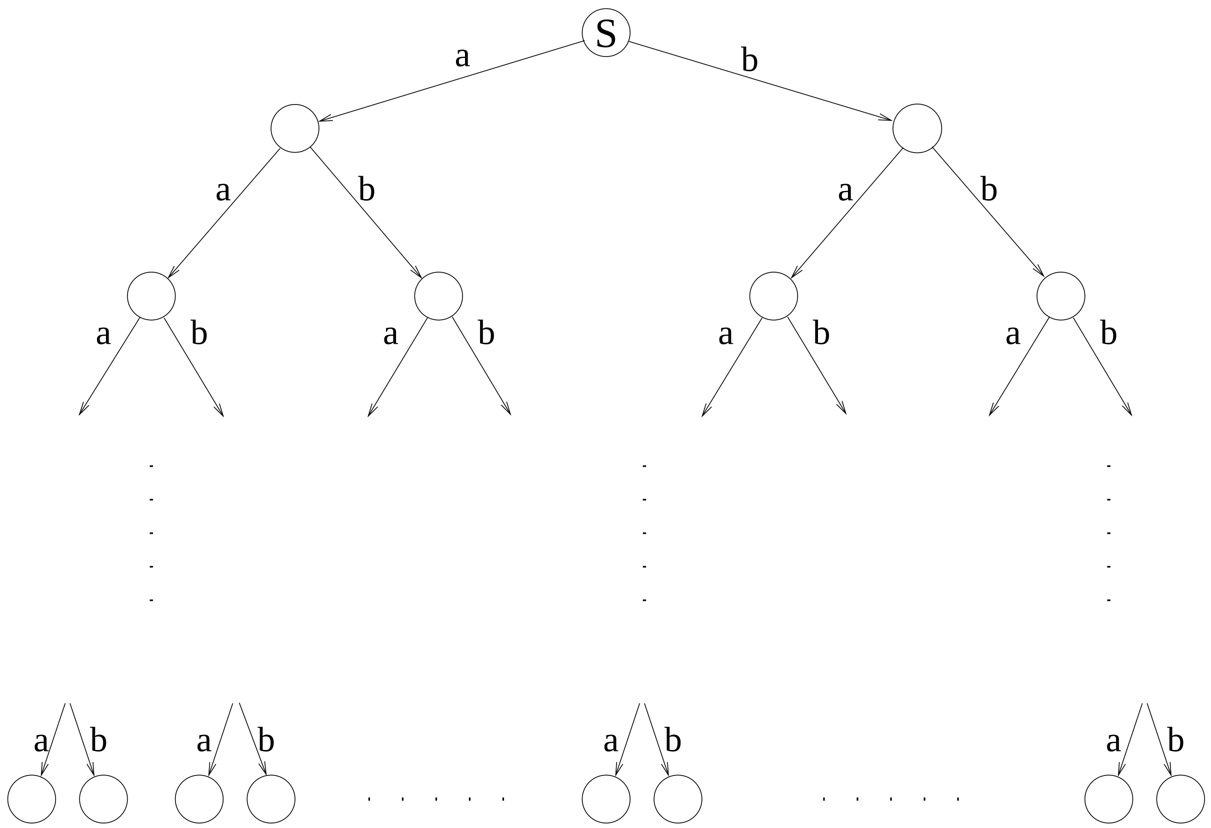
Nun verwenden wir n paarweise disjunkte Exemplare von a_n, um einen Akzeptor für L^n zu konstruieren. Hierbei verwenden wir die Assoziativität des Komplexproduktes und erhalten einen matrixartigen Akzeptor, der unten schematisch dargestellt wird.



Endliche Akzeptoren (26)

Offensichtlich gilt $|A| = n \cdot (2 \cdot n + 2)$. A ist nichtdeterministisch.

Der am einfachsten zu beschreibende deterministische Akzeptor B ist ein balancierter binärer Baum der Tiefe $2n$, dessen Wurzel der Startknoten ist.



Endliche Akzeptoren (27)

Markieren wir die Endknoten des Baumes als Terminalknoten, die zu den Wegen $w \in \mathcal{W}(G)$ mit $\mu(w) \in L_n$ gehören, dann erhalten wir einen Akzeptor für L_n.

Es stellt sich die Frage, ob einige der Kanten von B überflüssig sind.

Zwei Knoten P_1 und P_2 des Baumes kann man identifizieren, wenn die von P_1 und P_2 ausgehenden Wege bei gleicher Markierung zu gleichem Ergebnis führen: akzeptiert bzw. abgelehnt. Das ist dann der Fall, wenn in Punkt P_1 und P_2 schon entschieden ist, daß alle sich an P_1 und P_2 anschließenden Wege akzeptiert werden. Dieser Fall liegt genau dann vor, wenn der Weg von $w \in \mathcal{W}(S, P_1)$ an zwei „korrespondierenden" Stellen verschiedene Markierungen enthält.

Endliche Akzeptoren (28)

Diese Frage kann aber nicht entschieden werden, wenn es noch zu keiner Stelle des Weges eine korrespondierende Stelle gibt. Hieraus ergibt sich, daß man kein Punktepaar eines deterministischen Akzeptors identifizieren darf, das von S einen Abstand $\leq n$ besitzt. Hieraus folgt, daß für jeden vollständigen deterministischen Akzeptor $\tilde{a}$ die Abschätzung $|\tilde{a}| \geq 2^n$ gilt. Wir haben damit ein Beispiel gefunden, für das

$$|\tilde{a}| \geq 2^{\sqrt{|a|}}$$

gilt und zwar für **alle** deterministischen Akzeptoren von L_n, $n = 1, 2, \ldots$

Der minimale deterministische Akzeptor einer regulären Sprache L kann also wesentlich größer sein als ein nichtdeterministischer Akzeptor von L.