Technische Universität München
Fakultät für Informatik
Lehrstuhl Informatik XIV
Prof. Dr. Ernst W. Mayr
Dr. Werner Meixner

Sommersemester 2015
Lösungsblatt 2
4. Mai 2015

Theoretische Informatik

Hausaufgabe 1 (5 Punkte)

In der Vorlesung Diskrete Strukturen wird der Begriff des n-Tupels von Elementen eingeführt. Für eine beliebige Menge A wird gleichzeitig A^n als Bezeichnung für die Menge aller n-Tupel von Elementen aus A zusammen mit gleichbedeutenden Bezeichnungen

$$\underbrace{A \times A \times \ldots \times A}_{n-\text{fach}} \quad \text{oder} \quad A^{\times n}\,.$$

definiert. Mengentheoretisch gilt dabei stets $A^{n+1} \neq A^n \times A \neq A \times A^n$.

1. Sei Σ eine nichtleere Menge. Dann ist $(\Sigma^{\times 2}, \times_2)$ eine Algebra mit der 2-Tupelbildung $x \times_2 y := (x, y)$ als Operation. Zeigen Sie, dass die Operation $\times_2$ nicht assoziativ ist.

2. Sei Σ eine nichtleere endliche Menge. Dann ist die Konkatenation $\circ$ von Wörtern aus Σ^* eine assoziative Operation. Im Kontext der assoziativen Algebra $(\Sigma^*, \circ)$ wird das Produkt AB von Teilmengen $A, B \subseteq \Sigma^*$ definiert und die Potenzierung induktiv durch $A^{n+1} = AA^n$ eingeführt (siehe Vorlesung). Zeigen Sie, dass die folgende Gleichung für alle $A \subseteq \Sigma^*$ und $m, n \in \mathbb{N}_0$ gilt:
$$A^{m+n} = A^m A^n\,.$$
<u>Zu beachten:</u> Für eine beliebige Menge A ist i.A. keine Konkatenation definiert, wohl aber eine n-Tupelbildung.

3. Sei $R \subseteq [100] \times [100]$ eine binäre Relation über der Menge $[100] = \{1, 2, \ldots, 99, 100\}$ von natürlichen Zahlen mit $R = \{(x, y) \in [100] \times [100]\,;\, 3x = y\}$. Die Potenzierung R^n ist bekanntlich bezüglich der assoziativen Komposition $\circ$ von Relationen definiert.

   Berechnen Sie R^3 und $R^+ = \bigcup_{n \geq 1} R^n$ als endliche Listen.

Lösung

1. Wir geben $x, y, z \in \Sigma^{\times 2}$ an, so dass $l := x \times_2 (y \times_2 z) \neq (x \times_2 y) \times_2 z =: r$ gilt.

   Da $\Sigma \neq \emptyset$, kann somit $a \in \Sigma$ angenommen werden.

   Wir setzen $x = a$, $y = (a, a)$, $z = a$. Offenbar gilt $x, y, z \in \Sigma^{\times 2}$, da $\Sigma^{\times 2}$ die Menge Σ als Teilmenge enthält und abgeschlossen ist unter 2-Tupelbildung (siehe Blatt1, HA1).

   Dann gelten $r = (a, ((a, a), a))$ und $l = ((a, (a, a)), a)$. Nun folgt $r \neq l$, da die ersten Komponenten von r_1 bzw. l_2 von r bzw. l nicht gleich sind. Es gilt $r_1 = a \neq (a, (a, a)) = l_1$. W.z.b.w. (2P)

2. Zunächst folgt aus der Assoziativität der Konkatenation, dass für alle $A, B, C \subseteq \Sigma^*$ gilt $A(BC) = (AB)C$.

   Wir beweisen induktiv für alle $m \in \mathbb{N}_0$ die folgende Aussage $E(m)$.

   $$E(m): \qquad \forall n \in \mathbb{N}_0: \; A^{m+n} = A^m A^n$$

   <u>$m = 0$:</u> Für alle $n \in \mathbb{N}_0$ gilt $A^{0+n} = A^n = \{\epsilon\}A^n = A^0 A^n$.

   <u>Induktionsschritt $m \Rightarrow m+1$:</u>

   Sei $m \in \mathbb{N}_0$.

   Annahme: Es gelte $E(m)$. Wir zeigen, dass dann $E(m+1)$ gilt.

   Für alle $n \in \mathbb{N}_0$ gilt

   $$\begin{aligned} A^{(m+1)+n} &= AA^{m+n} && \text{(nach Definition)} \\ &= A(A^m A^n) && \text{(Ind. Annahme)} \\ &= (AA^m)A^n && \text{(Assoziativität)} \\ &= A^{m+1}A^n . && \text{(nach Definition)} \end{aligned}$$

   W.z.b.w. (2P)

3. Es gilt $R^3 = (R \circ R) \circ R = \{(x, y) \in [100] \times [100]\, ;\, 27x = y\}$.

   Damit gilt $R^3 = \{(1, 27), (2, 54), (3, 81)\}$.

   Wegen $R^n = \emptyset$ für alle $n \geq 5$ gilt

   $$R^+ = R^1 \cup R^2 \cup R^3 \cup R^4$$

   mit

   $$\begin{aligned} R^1 &= \{(1,3), (2,6), (3,9), \ldots \quad , (33,99)\}, \\ R^2 &= \{(1,9), (2,18), (3,27), \ldots, (11,99)\}, \\ R^3 &= \{(1,27), (2,54), (3,81)\}, \\ R^4 &= \{(1,81)\}. \end{aligned}$$

   (1P)

Hausaufgabe 2 (5 Punkte)

Sei Σ ein Alphabet. Zeigen Sie für alle formalen Sprachen A über Σ die folgenden Aussagen.

1. $A^* = A^+ \Leftrightarrow \epsilon \in A$. (Beachten Sie: $A^+ = AA^*$.)
2. $AA \subseteq A \;\Leftrightarrow\; A = A^+$.
3. Zeigen Sie für alle $A \neq \emptyset$: $\quad A \subseteq AA \;\Leftrightarrow\; \epsilon \in A$.

<u>Bemerkung:</u> In algebraischer Sprechweise heißt A *abgeschlossen* bezüglich der Konkatenation $\circ$, falls $AA \subseteq A$ gilt, und A ist in diesem Fall eine Unterhalbgruppe von $(\Sigma^*, \circ)$. Entsprechend ist A^+ die *von A erzeugte Halbgruppe* (oder Unterhalbgruppe). Falls $AA \subseteq A$ und $\epsilon \in A$ gelten, heißt A ein *Untermonoid* von $(\Sigma^*, \circ)$. A^* ist das *von A erzeugte Monoid* (oder Untermonoid).

Lösung

1. Beweis in zwei Richtungen:

   (a) Nehmen wir an, dass $A^* = A^+$ gilt. Da $\epsilon \in A^*$ gilt, ist auch $\epsilon \in A^+$. Mit $A^+ = AA^*$ folgt, dass es Wörter $u \in A$ und $v \in A^*$ geben muss, so dass $\epsilon = uv$. Dies gilt allerdings nur für $u = \epsilon$ (und $v = \epsilon$), weshalb $\epsilon \in A$ folgt.

   (b) Gelte $\epsilon \in A$. Dann gilt auch $\epsilon \in A^+$, denn $A^+ = AA^*$ und $\epsilon \in A^*$. Somit folgt:

$$A^* = \bigcup_{n \geq 0} A^n = A^0 \cup \bigcup_{n \geq 1} A^n = \{\epsilon\} \cup A^+ = A^+ \,.$$

(2P)

2. Beweis in zwei Richtungen:

   (a) Aus $AA \subseteq A$ folgt mit VA1(ii), Blatt 1 sofort $AAA \subseteq AA \subseteq A$, mithin $A^n \subseteq A$. Es folgt

$$A^+ = \bigcup_{n \geq 1} A^n \subseteq A \quad \text{und} \quad A \subseteq A^+ \,.$$

   (b) Nehmen wir an, dass $A = A^+$ gilt. Dann folgt:

$$AA = AA^+ \subseteq A^+ = A \,.$$

(2P)

3. Für die Konkatenation $\circ$ gilt die Gleichung $|u \circ v| = |u| + |v|$.
   Wir beweisen die Äquivalenz durch eine Kontraposition und eine Implikation.

   (a) Annahme $\epsilon \notin A$: Dann gilt für alle $u, v \in A$: $|u \circ v| = |u| + |v| > |u|$. Sei u_0 ein Element aus A mit minimaler Länge in A. Dann gilt für alle $x \in AA$ die Zerlegung $x = u \circ v$ mit gewissen $u, v \in A$, mithin $|x| = |u \circ v| = |u| + |v| \geq 2|u_0| > 0$. Es folgt $x \neq u_0$, d.h. $u_0 \notin AA$, mithin $A \not\subseteq AA$.

   (b) Annahme $\epsilon \in A$ und $x \in A$: Dann gilt $x = \epsilon \circ x \in AA$.

(1P)

Hausaufgabe 3 (5 Punkte)

Zeigen oder widerlegen Sie: Seien $u, v \in \Sigma^*$ Wörter mit $u \neq \epsilon$, $v \neq \epsilon$ und $uv = vu$. Dann existiert ein $z \in \Sigma^*$ mit $u = z^m$ und $v = z^n$ für gewisse $m, n \in \mathbb{N}$.
Hinweis: Verwenden Sie die Notation w_i, um den i-ten Buchstaben eines Wortes w zu bezeichnen. Dabei bezeichnet w_1 den ersten Buchstaben.

Lösung

Wir zeigen die Aussage per Induktion über $|uv|$ und nehmen dazu an, dass die Aussage bereits für alle Wörter $u', v' \in \Sigma^*$ mit $|u'v'| < |uv|$ gilt.
Für den Induktionsschritt können wir ohne Beschränkung der Allgemeinheit annehmen, dass $|u| \leq |v|$ gilt (ansonsten vertauschen wir u und v).
Zunächst einmal gilt: $u_i = (uv)_i = (vu)_i = v_i$ für alle $1 \leq i \leq |u|$. Haben wir $|u| = |v|$, so folgt damit $u = v$, und die Aussage gilt mit $z = u = v$ und $m = n = 1$.
Andernfalls gibt es ein $w \neq \epsilon$ mit $v = uw$ und es gilt: $uuw = uv = vu = uwu$ und damit $uw = wu$. Nach Induktionshypothese existieren also $m, k \in \mathbb{N}$ und ein Wort $z \in \Sigma^*$, so dass $z^m = u$ und $z^k = w$. Dann ist aber $v = uw = z^m z^k = z^{m+k}$, und die Aussage gilt mit $n = m + k$.

(5P)

Hausaufgabe 4 (5 Punkte)

Sei $\Sigma = \{(,)\}$ der Zeichenvorrat mit einer öffnenden und einer schließenden Klammer. Für $w \in \Sigma^*$ definieren wir $|w|_($ bzw. $|w|_)$ als die Anzahl der in w enthaltenen öffnenden bzw. schließenden Klammern. u ist ein Anfangsteilwort (Praefix) von w, falls es ein Wort v gibt, so dass $w = uv$ gilt. Wir nennen ein nichtleeres Wort $w \in \Sigma^*$ positiv, falls $|u|_(> |u|_)$ für alle nichtleeren Anfangsteilwörter u von w gilt.

Bestimmen Sie die Anzahl der positiven Wörter über Σ der Länge $n \in \mathbb{N}$!

Hinweis: Benutzen Sie die Formel zur Lösung des Ballot-Problems aus der Vorlesung Diskrete Strukturen (WS 12/13) wie folgt.
Ballot-Problem: Bei einer Wahl erhält Kandidat A a Stimmen und Kandidat B b Stimmen, mit $a > b \geq 0$. Die Stimmzettel werden sequentiell ausgezählt. Wie viele Zählfolgen gibt es, so dass A nach jedem Schritt in Führung ist?
Lösung: Die gesuchte Anzahl ist $\frac{a-b}{a+b}\binom{a+b}{a}$, oder a.a., $\frac{a-b}{a+b}\binom{a+b}{b}$.

Lösung

Ballot-Problem: Bei einer Wahl erhält Kandidat A a Stimmen und Kandidat B b Stimmen, mit $a > b \geq 0$. Die Stimmzettel werden sequentiell ausgezählt. Wie viele Zählfolgen gibt es, so dass A nach jedem Schritt in Führung ist?
Lösung: Die gesuchte Anzahl ist $\frac{a-b}{a+b}\binom{a+b}{a}$, oder a.a., $\frac{a-b}{a+b}\binom{a+b}{b}$.

Anwendung:
Für ein positives Wort $w \in \Sigma^*$ gilt $|w| = |w_(| + |w_)| = n$ mit $a := |w_(| > |w_)| =: b$. Die Anzahl $A_{b,n}$ der positiven Wörter mit b schließenden Klammern ist also

$$A_{b,n} = \frac{n-2b}{n}\binom{n}{b}.$$

Die Anzahl $A_{p,n}$ der positiven Wörter der Länge n ist demnach

$$A_{p,n} = \sum_{b=0}^{\lfloor \frac{n-1}{2} \rfloor} \frac{n-2b}{n}\binom{n}{b}.$$

(5P)

Hinweis: *Die Vorbereitungsaufgaben bereiten die Tutoraufgaben vor und werden in der Zentralübung unterstützt. Tutoraufgaben werden in den Übungsgruppen bearbeitet. Hausaufgaben sollen selbstständig bearbeitet und zur Korrektur und Bewertung abgegeben werden.*

Vorbereitung 1

Eine Grammatik G sei gegeben in BNF-Form durch

$$S \to a\,S\,d\,d, \qquad\qquad S \to \{b\} \mid \{c\}.$$

Geben Sie G als kontextfreie Grammatik $G = (V, \{a, b, c, d\}, P, S)$ an.

Lösung

Wir lösen die BNF-Form nach Vorlesung wie folgt auf.

Der Ausdruck $S \to \{b\} \mid \{c\}$ steht für die beiden Ausdrücke $S \to \{b\}$ und $S \to \{c\}$.
Der Ausdruck $S \to \{b\}$ steht für die Menge von Regeln $\{S \to \epsilon, S \to X, X \to b, X \to bX\}$.
Der Ausdruck $S \to \{c\}$ steht für die Menge von Regeln $\{S \to \epsilon, S \to Y, Y \to c, Y \to cY\}$.
Dabei sind X, Y Variable, die nicht in G vorkommen.
Es bleibt noch die Umbezeichnung von S nach T, um die Monotoniebedingung zu sichern.
Deswegen muss auch $T \to add$ hinzugefügt werden.

Wir fassen zusammen.

$$\begin{aligned} V &:= \{S, X, Y, T\}\,, \\ P &:= \{\,T \to aTdd\,, T \to add\,, \\ &\quad\; T \to X\,,\, X \to b\,,\, X \to bX\,, \\ &\quad\; T \to Y\,,\, Y \to c\,,\, Y \to cY\,, \\ &\quad\; S \to \epsilon\,,\, S \to T\,\}\,. \end{aligned}$$

Vorbereitung 2

Gegeben sind folgende Grammatiken:

$$G_1 := (\{S\},\ \{\mathtt{a}, \mathtt{b}, \mathtt{+}, \mathtt{(}, \mathtt{)}\},\ \{S \to \mathtt{a},\ S \to \mathtt{b},\ S \to S\mathtt{+}S,\ S \to \mathtt{(}S\mathtt{)}\},\ S),$$
$$G_2 := (\{S\},\ \{\mathtt{a}, \mathtt{b}, \mathtt{+}, \mathtt{(}, \mathtt{)}\},\ \{S \to \mathtt{a},\ S \to \mathtt{b},\ S \to \mathtt{a+}S,\ S \to \mathtt{b+}S,\ S \to \mathtt{(}S\mathtt{)}\},\ S).$$

1. Ordnen Sie die Grammatiken in die Chomsky-Hierarchie ein.

2. Geben Sie jeweils einen Ableitungsbaum für das Wort `a+(b+a)` an.

3. Gilt $L(G_1) = L(G_2)$?

Lösung

1. Die Frage der Einordnung einer Grammatik in die Chomsky-Hierarchie ist erst dann vollständig beantwortet, wenn gesagt wird, welchen Typ sie besitzt und gleichzeitig gesagt wird, dass sie den nächsthöheren Typ nicht besitzt. Wir zeigen, dass beide Grammatiken vom Typ 2 und nicht vom Typ 3 sind.

   Damit eine Grammatik vom Typ 2 ist, genügt es nicht, dass alle Regeln die Form $\alpha \to \beta$ mit $\alpha \in V$ haben. Es muss auch die Monotoniebedingung erfüllt sein, d. h.

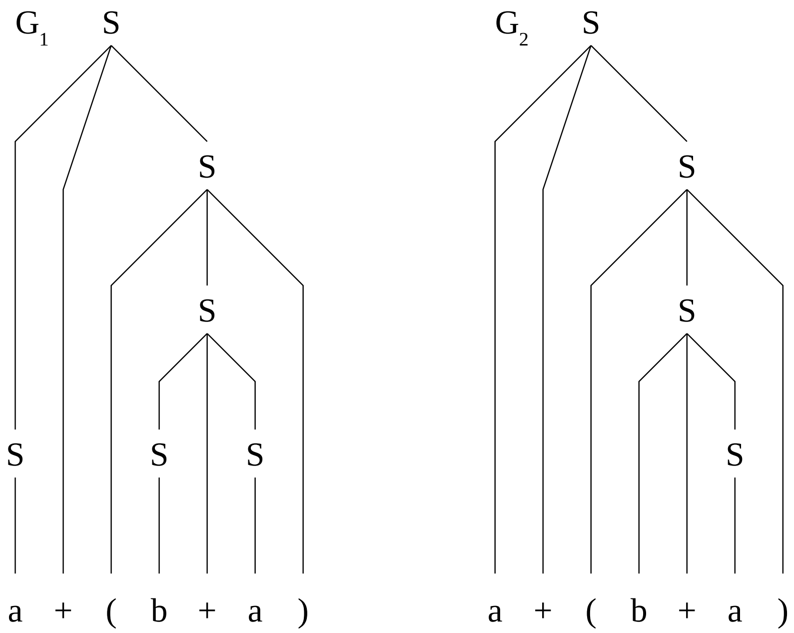


Abbildung 1: Ableitungsbäume für `a+(b+a)`

$|\alpha| \leq |\beta|$ mit der leichten Abschwächung, dass diese Ungleichung nur für $\alpha \neq S$ gefordert wird und im Fall $\beta = \epsilon$ und $\alpha = S$ gleichzeitig die Bedingung gelten muss, dass S nie auf der rechten Seite einer Produktion vorkommt.

Da für alle Regeln $\alpha \rightarrow \beta$ beider Grammatiken $\beta \neq \epsilon$ und außerdem $\alpha \in V$ gilt, folgt trivialerweise sogar die strikte Monotonie der Grammatiken, d. h. $|\alpha| \leq |\beta|$ für alle Regeln. Deshalb sind wegen $\alpha \in V$ beide Grammatiken vom Typ 2 oder kontextfrei.

Beide Grammatiken sind nicht regulär, d. h. nicht vom Typ 3. Als Beleg dafür kann die in beiden Grammatiken enthaltene Produktion $S \rightarrow \texttt{(}S\texttt{)}$ dienen, die weder rechtslinear noch linkslinear ist.

2. Ableitungsbaum siehe Abbildung 1.

3. Zunächst gilt für $w = \texttt{(a+a)+a}$, dass $w \in L(G_1)$ wegen

$$S \underset{G_1}{\rightarrow} S\texttt{+}S \underset{G_1}{\rightarrow} \texttt{(}S\texttt{)+}S \underset{G_1}{\rightarrow}^{*} \texttt{(}S\texttt{+}S\texttt{)+}S \underset{G_1}{\rightarrow}^{*} \texttt{(a+a)+a}\,.$$

Andererseits gilt $w \notin L(G_2)$. Zum Beweis dieser Aussage nehmen wir $S \underset{G_2}{\rightarrow}^{*} \texttt{(a+a)+a}$ an. Im allerersten Ersetzungsschritt kann keine Produktion angewandt werden, die ein Terminalzeichen `a` oder `b` am Anfang der hergeleiteten Satzform produziert. D. h., es kann nur die Produktion $S \rightarrow \texttt{(}S\texttt{)}$ zur Anwendung kommen, also gilt

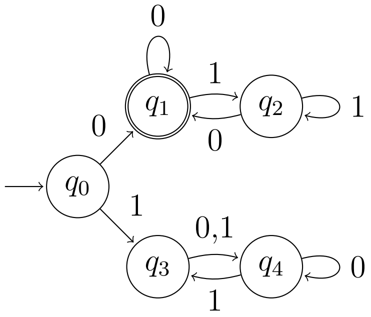
$$S \underset{G_2}{\rightarrow} \texttt{(}S\texttt{)} \underset{G_2}{\rightarrow}^{*} \texttt{(a+a)+a}\,.$$

Dies ist aber sicher falsch, denn die schließende Klammer kann in keinem Ersetzungsschritt entfernt werden, mithin gilt $\texttt{(a+a)+a} \notin L(G_2)$.

Es folgt $L(G_1) \neq L(G_2)$.

Vorbereitung 3

Wir betrachten einen endlichen deterministischen Automat $A = (Q, \Sigma, \delta, q_0, F)$, der durch die folgende Grafik gegeben ist.



1. Übersetzen Sie die Grafik in eine extensionale Mengenschreibweise (Darstellung durch Auflistung) für Q, Σ, δ und F.
2. Bestimmen Sie $\delta(\delta(q_1, 0), 1)$ und $\hat{\delta}(q_0, 10)$!
3. Geben Sie ein möglichst einfaches Kriterium an, mit dem man entscheiden kann, ob ein Wort $w \in \Sigma^*$ von A akzeptiert wird.

Lösung

1. $$\begin{aligned} Q &= \{q_0, q_1, q_2, q_3, q_4\}\,, \\ \Sigma &= \{0, 1\}\,, \\ \delta &= \{((q_0, 0), q_1),\ ((q_0, 1), q_3), \\ &\quad ((q_1, 0), q_1),\ ((q_1, 1), q_2), \\ &\quad ((q_2, 0), q_1),\ ((q_2, 1), q_2), \\ &\quad ((q_3, 0), q_4),\ ((q_3, 1), q_4), \\ &\quad ((q_4, 0), q_4),\ ((q_4, 1), q_3)\}\,, \\ F &= \{q_1\}\,. \end{aligned}$$

   Startzustand ist q_0.

2. $\delta(\delta(q_1, 0), 1) = \delta(q_1, 1) = q_2\,,$

   $\hat{\delta}(q_0, 10) = \hat{\delta}(\delta(q_0, 1), 0) = \hat{\delta}(q_3, 0) = \hat{\delta}(\delta(q_3, 0), \epsilon) = \delta(q_3, 0) = q_4\,.$

3. Wörter, die mit 1 beginnen, werden nicht akzeptiert, weil in dem betreffenden Zweig kein Endzustand erreichbar ist. Andererseits endet offenbar jedes Wort, das in den Endzustand führt, mit dem Zeichen 0.

   Wir vermuten, dass alle diejenigen Wörter akzeptiert werden, die mit 0 beginnen und mit 0 enden.

   Sei $w = 0$. Offenbar beginnt w mit 0 und endet gleichzeitig mit 0. Und A akzeptiert w.

   Nun sei $w = 0u0$ ein Wort mit $u \in \Sigma^*$. Die Eingabe von 0 führt in den Zustand q_1. Von q_1 aus kommt man bei beliebiger Eingabe nur entweder nach q_1 oder q_2. Gleich wo man gelandet ist, führt die Eingabe von 0 wieder in den Endzustand q_1. D. h., dass w akzeptiert wird.

Vorbereitung 4

Geben Sie jeweils einen endlichen Automat (als Graph und Übergangsrelation) an, der über dem Alphabet $\Sigma = \{0, 1\}$ folgende Sprache akzeptiert:

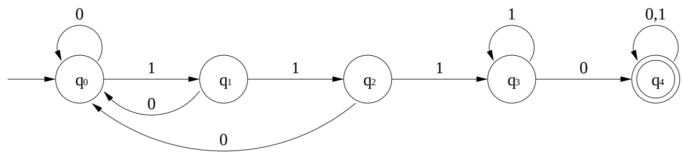
1. Die Menge aller Wörter, die das Teilwort 1110 enthalten.

2. Die Menge aller Wörter, bei denen die Anzahl der Einsen durch 3 teilbar ist.

3. Die Menge aller Wörter, die mit 10 beginnen und auf 01 enden.

Lösung

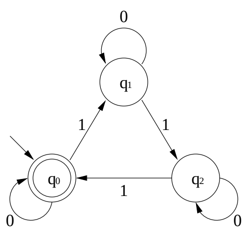
1. Die Menge aller Wörter, die das Teilwort 1110 enthalten.



Übergangsrelation:

q_i	$\delta(q_i, 0)$	$\delta(q_i, 1)$
q_0	q_0	q_1
q_1	q_0	q_2
q_2	q_0	q_3
q_3	q_4	q_3
q_4	q_4	q_4

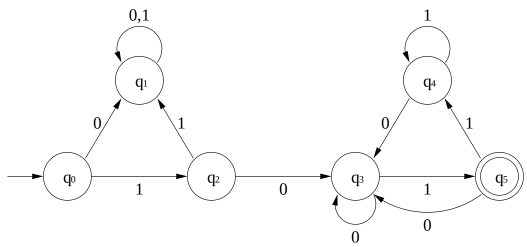
2. Die Menge aller Wörter, bei denen die Anzahl der Einsen durch 3 teilbar ist.



Übergangsrelation:

q_i	$\delta(q_i, 0)$	$\delta(q_i, 1)$
q_0	q_0	q_1
q_1	q_1	q_2
q_2	q_2	q_0

3. Die Menge aller Wörter, die mit 10 beginnen und auf 01 enden.



Übergangsrelation:

q_i	$\delta(q_i, 0)$	$\delta(q_i, 1)$
q_0	q_1	q_2
q_1	q_1	q_1
q_2	q_3	q_1
q_3	q_3	q_5
q_4	q_3	q_4
q_5	q_3	q_4

Tutoraufgabe 1

Wir beziehen uns auf die in der Vorbereitungsaufgabe 2 definierten Grammatiken G_1 und G_2.

1. Zeigen Sie: Für alle $w_1, w_2 \in L(G_1)$ ist auch $\texttt{(}w_1\texttt{+}w_2\texttt{)} \in L(G_1)$.

   Gilt diese Aussage auch für G_2?

2. Sind die Grammatiken G_1 bzw. G_2 eindeutig?

Lösung

1. (a) Wenn $w_1, w_2 \in L(G_1)$, so gilt $S \underset{G_1}{\rightarrow}^* w_1$ und $S \underset{G_1}{\rightarrow}^* w_2$. Es folgt

$$S \underset{G_1}{\rightarrow} \texttt{(}S\texttt{)} \underset{G_1}{\rightarrow} \texttt{(}S\texttt{+}S\texttt{)} \underset{G_1}{\rightarrow}^* \texttt{(}w_1\texttt{+}S\texttt{)} \underset{G_1}{\rightarrow}^* \texttt{(}w_1\texttt{+}w_2\texttt{)}\,.$$

   Man beachte dabei, dass die Bildung der Halbgruppenhülle der Ableitungsrelation $\underset{G_1}{\rightarrow}^*$ nicht aus $\underset{G_1}{\rightarrow}^*$ herausführt.

   (b) Für $L(G_2)$ gilt die Aussage nicht, d. h. es existieren $w_1, w_2 \in L(G_2)$, so dass $\texttt{(}w_1\texttt{+}w_2\texttt{)} \notin L(G_2)$.

   Zum Beweis dieser Aussage erinnern wir uns an den 3. Teil der Vorbereitungsaufgabe 2, wo gezeigt wurde, dass $\texttt{(a+a)+a} \notin L(G_2)$. Falls wir nun zeigen können, dass sogar auch $\texttt{((a+a)+a)} \notin L(G_2)$ gilt, dann brauchen wir nur noch $w_1 = \texttt{(a+a)}$ und $w_2 = \texttt{a}$ zu setzen. Da offenbar $w_1, w_2 \in L(G_2)$ und $\texttt{(}w_1\texttt{+}w_2\texttt{)} \notin L(G_2)$, haben wir dann die gewünschte Aussage bewiesen.

   Wir müssen zunächst beachten, dass der einfache Schluss von $\texttt{(a+a)+a} \notin L(G_2)$ auf $\texttt{((a+a)+a)} \notin L(G_2)$ i. A. nicht zulässig ist, d. h. für die spezielle Grammatik G_2 erst bewiesen werden muss. Wir beweisen den Schluss in der indirekten Form

$$\texttt{((a+a)+a)} \in L(G_2) \quad \Longrightarrow \quad \texttt{(a+a)+a} \in L(G_2)$$

   und nehmen an $S \underset{G_2}{\rightarrow}^* \texttt{((a+a)+a)}$. Es folgt $S \underset{G_2}{\rightarrow} \texttt{(}S\texttt{)} \underset{G_2}{\rightarrow}^* \texttt{((a+a)+a)}$, denn die Anwendung jeder anderen Produktion im ersten Schritt würde links ein anderes Terminalzeichen produzieren als die öffnende Klammer. Für Terminalzeichen x nimmt die Zerlegungseigenschaft der Ableitungsrelation in kontextfreien Sprachen die besondere Form an

$$xy \underset{G_2}{\rightarrow}^* xy' \quad \Longrightarrow \quad y \underset{G_2}{\rightarrow}^* y'\,.$$

Die Zerlegungseigenschaft in verschiedenen Varianten wird in Vorbereitungsaufgabe 2, Blatt 1 behandelt. Sie ist eine zentrale und oft gebrauchte Eigenschaft von kontextfreien Sprachen.

Aus

$$(S) \underset{G_2}{\rightarrow}^* \texttt{((a+a)+a)}$$

folgt deshalb

$$S) \underset{G_2}{\rightarrow}^* \texttt{(a+a)+a)}$$

und daraus analog für die schließende Klammer auf der rechten Seite

$$S \underset{G_2}{\rightarrow}^* \texttt{(a+a)+a}\,.$$

Dies steht im Widerspruch zu der in der Vorbereitungsaufgabe 2 bewiesenen Aussage $\texttt{(a+a)+a} \notin L(G_2)$. Also gilt $\texttt{((a+a)+a)} \notin L(G_2)$.

2. (a) Die Variablen in Wortformen von kontextfreien Grammatiken können grundsätzlich in beliebiger Reihenfolge verarbeitet, d. h. durch Anwendung von Produktionen ersetzt werden. Ableitungen eines Terminalwortes w, die sich nur in der Reihenfolge der Verarbeitung der Variablen unterscheiden, sind voneinander nicht wesentlich verschieden. Beim Vergleich von Ableitungen verwendet man deshalb stets Linksableitungen, d. h. Ableitungen $\alpha_0, \alpha_1, \ldots, \alpha_n$, in denen stets die am weitesten links stehende Variable in α_i im nächsten Anwendungsschritt $i+1$ ersetzt wird.

   Das Wort $w = \texttt{a+a+a}$ besitzt die folgenden beiden Linksableitungen in G_1

$$\underline{S} \underset{G_1}{\rightarrow} \underline{S}\texttt{+}S \underset{G_1}{\rightarrow} \texttt{a+}\underline{S} \underset{G_1}{\rightarrow} \texttt{a+}\underline{S}\texttt{+}S \underset{G_1}{\rightarrow} \texttt{a+a+}\underline{S} \underset{G_1}{\rightarrow} \texttt{a+a+a}$$
$$\underline{S} \underset{G_1}{\rightarrow} \underline{S}\texttt{+}S \underset{G_1}{\rightarrow} \underline{S}\texttt{+}S\texttt{+}S \underset{G_1}{\rightarrow} \texttt{a+}\underline{S}\texttt{+}S \underset{G_1}{\rightarrow} \texttt{a+a+}\underline{S} \underset{G_1}{\rightarrow} \texttt{a+a+a}.$$

   Die in einem Ersetzungsschritt verarbeitete Variable ist hier unterstrichen. Wie man sieht, steht die unterstrichene Variable in beiden Ableitungen stets links von allen anderen Variablen.

   Offensichtlich wurden im 2. Ersetzungsschritt in den beiden Ableitungen unterschiedliche Ersetzungen der linken Variablen S vorgenommen. Die Ableitungen sind also sicherlich verschieden, obwohl beide Ableitungen letztendlich das gleiche Terminalwort $w = \texttt{a+a+a}$ ableiten.

   Daraus folgt, dass G_1 mehrdeutig ist.

   (b) Hingegen ist G_2 eindeutig. Wir zeigen dies, indem wir die Annahme zum Widerspruch führen, dass es ein Wort aus $L(G_2)$ gibt, das 2 verschiedene Linksableitungen besitzt.

   Sei also $w \in L(G_2)$ und seien $S = \alpha_0, \alpha_1, \ldots, \alpha_m = w$ und $S = \beta_0, \beta_1, \ldots, \beta_n = w$ zwei nicht identische Linksableitungen α, β. Wir nehmen $m \le n$ o. B. d. A. an. Sei i der Index, so dass sich die Folgen erstmalig in α_i und β_i unterscheiden, d. h. $\alpha_i \ne \beta_i$ und $\alpha_j = \beta_j$ für alle $j < i$. Einen solches i gibt es, weil sonst die Gleichheit beider Ableitungen folgen würde. Dann gibt es $x, y, A, \gamma', \gamma''$ mit $\gamma' \ne \gamma''$, so dass

$$\begin{array}{lllllllll} S & \underset{G_2}{\rightarrow}^* & \alpha_{i-1} & = & xAy & \underset{G_2}{\rightarrow} & x\gamma' y & = \alpha_i & \underset{G_2}{\rightarrow}^* & w\,, \\ S & \underset{G_2}{\rightarrow}^* & \beta_{i-1} & = & xAy & \underset{G_2}{\rightarrow} & x\gamma'' y & = \beta_i & \underset{G_2}{\rightarrow}^* & w\,, \end{array}$$

wobei x keine Variable enthalte, weil die Variable A die am weitesten links stehende Variable sein soll, $y \in (V \cup \Sigma)^*$, $\gamma', \gamma'' \in (V \cup \Sigma)^+$.

Es folgt, dass x ein Präfix ist von w, d. h. $w = xu$. Nun aber bestimmt die Gestalt von u eindeutig die auf A anzuwendende Regel und es folgt $\gamma' = \gamma''$, im Widerspruch zu $\gamma' \neq \gamma''$. Denn entweder ist u einelementig, was die Anwendung genau einer der Terminalregeln erfordert, oder u ist mehrelementig. Dann aber bestimmt wiederum der erste Buchstabe von u genau die anzuwendende Regel von denjenigen Regeln, die mehrere Zeichen produzieren.

Tutoraufgabe 2

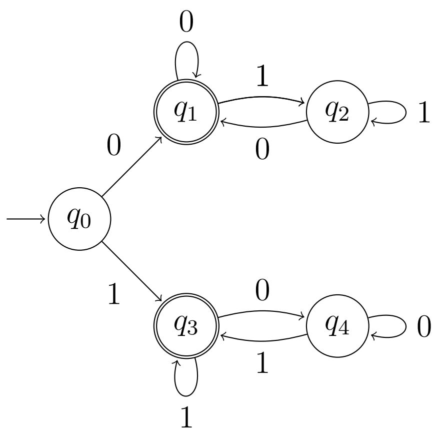
Wir betrachten die Sprache L aller Wörter über dem Alphabet $\Sigma = \{0, 1\}$, die entweder mit 1 beginnen und gleichzeitig mit 1 enden oder die mit 0 beginnen und gleichzeitig mit 0 enden.

1. Geben Sie einen deterministischen endlichen Automaten (DFA) an, der L akzeptiert, und zeigen Sie, dass es unendlich viele DFA gibt, die L akzeptieren.

2. Geben Sie einen nichtdeterministischen endlichen Automaten (NFA) mit höchstens 4 Zuständen an, der L akzeptiert.

Lösung

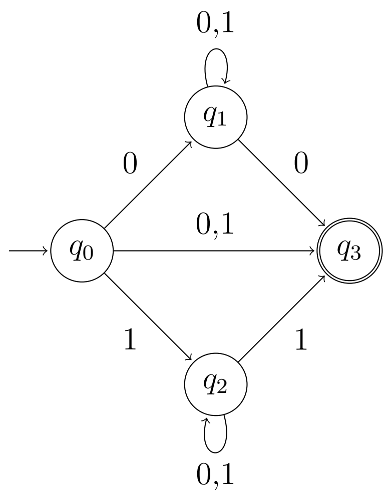
1. Entsprechender (minimaler) deterministischer Automat:



q_i	$\delta(q_i, 0)$	$\delta(q_i, 1)$
q_0	q_1	q_3
q_1	q_1	q_2
q_2	q_1	q_2
q_3	q_4	q_3
q_4	q_4	q_3

Automaten A und B unterscheiden sich bereits dann, wenn die Menge der Zustände nicht in beiden Automaten gleich ist. Deshalb kann man für jeden Automaten A beliebig viele Automaten B angeben, so dass die akzeptierten Sprachen alle gleich sind, indem man Zustände hinzufügt, die vom Startzustand aus nicht erreichbar sind.

2. Nichtdeterministischer Automat, der L akzeptiert:



q_i	$\delta(q_i,0)$	$\delta(q_i,1)$
q_0	$\{q_1,q_3\}$	$\{q_2,q_3\}$
q_1	$\{q_1,q_3\}$	$\{q_1\}$
q_2	$\{q_2\}$	$\{q_2,q_3\}$
q_3	$\emptyset$	$\emptyset$

Tutoraufgabe 3

Sei $\Sigma = \{0,1,2,3\}$ die Zeichenmenge der Ziffern von 0 bis 3. Sei Q die Sprache der Zahldarstellungen zur Basis 4 ohne führende Nullen. $\#(x)$ sei die der Darstellung x zugeordnete ganze Zahl. (Beispiel: $0 \in Q$, $2013 \in Q$, $02013 \notin Q$. Es gilt $\#(2013_4) = 4 \cdot 4^4 + 2 \cdot 4^3 + 4^1 + 3 \cdot 4^0 = \#(135_{10})$.)

Sei $L = \{w \in Q\,;\ \#(w) \bmod 3 = 2\}$.

1. Konstruieren Sie einen deterministischen endlichen Automaten A, der L akzeptiert.

2. Beweisen Sie, dass A die Sprache L akzeptiert, d. h., dass $L(A) = L$ gilt.

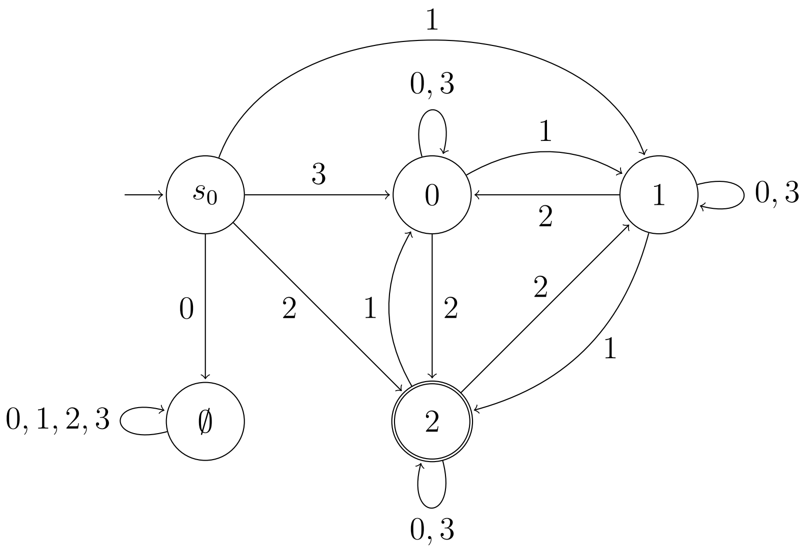
Lösung

Für $w \in \Sigma^*$ und $a \in \Sigma$ gilt $\#(wa) = 4 \cdot (\#w) + \#a$, mithin gilt $\#(wa) \bmod 3 = ((4 \bmod 3) \cdot (\#w \bmod 3) + \#a \bmod 3) \bmod 3 = (\#w \bmod 3 + \#a) \bmod 3)$.
Wenn wir $\#w = 0$ für $w = \epsilon$ setzen, dann gilt die Rekursion

$$\#(wa) \ \bmod 3 = \begin{cases} (0 + \#a) \ \bmod 3 & : \ \#w \ \bmod 3 = 0\,, \\ (1 + \#a) \ \bmod 3 & : \ \#w \ \bmod 3 = 1\,, \\ (2 + \#a) \ \bmod 3 & : \ \#w \ \bmod 3 = 2\,. \end{cases} \tag{1}$$

1. Die Konstruktion folgt Gleichung 1 und sorgt für die korrekte Erkennung und Verarbeitung von Q. Wir setzen $A = (\{s_0, 0, 1, 2, \emptyset\}, \Sigma, \delta, s_0, \{2\})$, so dass für alle $q \in \{0,1,2\}$ und $a \in \Sigma$
$$\delta(q,a) = (q + \#a) \ \bmod 3 \tag{2}$$

und damit folgendes Übergangsdiagramm gilt:



2. Dass führende Nullen nicht akzeptiert werden, ist klar. Auch die 0 wird korrekterweise nicht akzeptiert. Jede andere Eingabe eines einzelnen Zeichens führt in einen der Zustände 0, 1 oder 2, wobei die Zustände korrekt die Reste modulo 3 darstellen.

   Wir fassen die Zustände 0, 1 und 2 als entsprechende Zahlen 0, 1 bzw. 2 auf und stellen fest, dass für alle $w \in Q \setminus \{0\}$ mit $|w| = 1$ die folgende Gleichung gilt:

   $$\hat{\delta}(s_0, w) = (\#w) \mod 3\,. \tag{3}$$

   Wir beweisen Gleichung 3 für alle $w \in Q \setminus \{0\}$ mittels Induktion über die Länge von w.

   $\underline{|w| = 1}$: Wurde oben gezeigt.

   $\underline{n \to n{+}1 \text{ für alle } n \geq 1}$: Sei $n \geq 1$. Gleichung 3 gelte für alle w mit $|w| = n$. Dann folgt mit Hilfe von Gleichungen 1 und 2 zusammen mit der Induktionsannahme

   $$\begin{aligned} \hat{\delta}(s_0, wa) &= \delta(\hat{\delta}(s_0, w), a) \\ &= (\hat{\delta}(s_0, w) + \#a) \mod 3 \\ &= (\#w \mod 3 + \#a) \mod 3 \\ &= \#(wa) \mod 3\,. \end{aligned}$$

   Da 2 der einzige Endzustand ist, ist L die akzeptierte Sprache. Wzbw.