FGI-1 – Formale Grundlagen der Informatik I

Logik, Automaten und Formale Sprachen

Musterlösung 2 : ϵ-FA und Pumping-Lemma

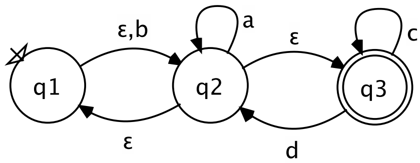
Variante für die Übungsgruppenleiter

Präsenzaufgabe 2.1: Wir festigen das Konzept der ϵ-Übergänge. Außerdem üben wir die Konstruktion, ϵ-Übergänge zu entfernen, ein. Letztere wird in den Hausaufgaben wieder aufgegriffen, sollte daher in der Übung i.w. verstanden sein.

1. Berechnen Sie die ϵ-Hülle, d.h. die Relation $R \subseteq Q \times Q$ mit

$$R = \{(q, q') \mid (q, \epsilon) \vdash^* (q', \epsilon)\}$$

für den folgenden ϵ-NFA.



Lösung: Wir können die Relation auch als Vereinigung darstellen: $R := \bigcup_{i \geq 0} R_i$, wobei R_i die Menge der Zustandspaare darstellt, die über genau i ϵ-Kanten verbunden werden, d.h.

$$R_i = \{(q, q') \mid (q, \epsilon) \vdash^i (q', \epsilon)\}$$

Wir lesen vom Zustandsdiagramm ab:

$$R_0 = \begin{array}{c|ccc} & q_1 & q_2 & q_3 \\ \hline q_1 & 1 & & \\ q_2 & & 1 & \\ q_3 & & & 1 \end{array} \qquad R_1 = \begin{array}{c|ccc} & q_1 & q_2 & q_3 \\ \hline q_1 & & 1 & \\ q_2 & 1 & & 1 \\ q_3 & & & \end{array} \qquad R_2 = \begin{array}{c|ccc} & q_1 & q_2 & q_3 \\ \hline q_1 & 1 & & 1 \\ q_2 & & 1 & \\ q_3 & & & \end{array}$$

Für alle höheren Ordnungen (d.h. für $i \geq 3$) existieren keine neuen Verbindungen mehr. Damit ergibt sich:

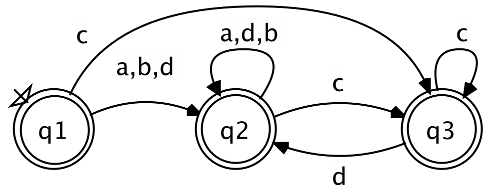
$$R := \bigcup_{i \geq 0} R_i = \bigcup_{i=0}^{2} R_i = \begin{array}{c|ccc} & q_1 & q_2 & q_3 \\ \hline q_1 & 1 & 1 & 1 \\ q_2 & 1 & 1 & 1 \\ q_3 & & & 1 \end{array}$$

Das Gleiche als Relation notiert ist dann:

$$R = \{(q_1, q_2), (q_2, q_3), (q_2, q_1), (q_1, q_3)\} \cup Id_Q$$

2. Konstruieren Sie für den obigen ϵ-FA einen äquivalenten ϵ-freien NFA.

Lösung: (Vgl. dazu die Definition in Satz 14.1.) Die Übergänge $q \xrightarrow{x} q''$ des äquivalenten ϵ-freien NFA ergeben sich, indem man zunächst im Orginalautomat mit beliebig vielen ϵ-Schritten von q zu einem q' und von dort mit x zu q'' gelangt. Die Endzustände ergeben sich, indem man „rückwärts", von den Endzuständen startend beliebig viele ϵ-Schritten läuft.



Präsenzaufgabe 2.2: Wir thematisieren das Pumping-Lemma. Dies bereitet auf die Hausaufgabe vor. Beim Beweis zu Teilaufgabe 1 ist insbesondere das Schematische an der Beweisfigur zu betonen, d.h. die Struktur, die bei allen Pumping-Lemma-Widerspruchsbeweisen wiederkehrend gleich ist.

Teilaufgaben 2 und 3 können auch sehr knapp abgehandelt werden.

1. Zeigen Sie mit Hilfe des Pumping-Lemmas, dass die Sprache $L = \{a^k b^{2k} \mid k \in \mathbb{N}\}$ nicht regulär ist.

   Lösung: *Pumping Lemma: Sei L eine reguläre Sprache. Dann existiert eine Zahl n, so dass für alle $z \in L$ mit $|z| \geq n$ stets eine Zerlegung $z = uvw$ existiert, so dass gilt:*
   (i) $|uv| \leq n$
   (ii) $|v| \geq 1$
   (iii) $\forall i \in \mathbb{N} : uv^i w \in L$

   Angenommen L wäre regulär. Wähle $z = a^n b^{2n}$ für die Zahl n des PL. Da $|z| \geq n$, muss es eine Zerlegung $z = uvw$ mit obigen Eigenschaften geben. Dann muss $uv \in \{a\}^*$ sein, d.h. $v = a^l$ für ein $l > 0$, denn nach (i) ist $|uv| \leq n$. Nach dem PL müsste dann das Wort $uv^0w = a^{n-l}b^{2n}$ in L sein, was aber nicht der Fall ist. Widerspruch.

2. Zeigen Sie, dass jede endliche Menge regulär ist.

   Lösung: Sei $L = \{w_1, \ldots, w_n\} \subseteq \Sigma^*$.

   Mit einem GFA (folgt noch in der Vorlesung) können wir diese Menge akzeptieren, wenn wir nur zwei Zustände q_0 und q_1 haben und für jedes Wort $w_i \in L$ eine mit w_i beschriftete Kante von q_0 nach q_1 haben, wobei q_0 der Start- und q_1 der einzige Endzustand ist.

   Alternativ können wir die Menge auch durch einen NFA akzeptieren. Wir definieren die Zustandsmenge als die Menge aller Suffixe der Worte aus L.

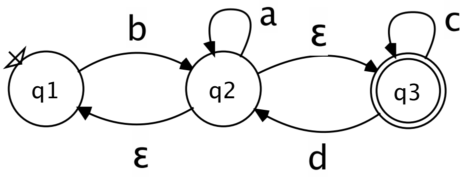
3. Die Sprache $L = \{a, ab, ac\}$ ist regulär. Zeigen Sie, dass das Pumping-Lemmas auch auf diese Sprache L zutrifft.

   Lösung: Beachte: Das PL sagt nicht, dass jede reguläre Menge unendlich groß wäre. Dies könnte man annehmen, da eine Eigenschaft des PL besagt, dass $\{u\}\{v\}^*\{w\} \subseteq L$ gilt, d.h. dass eine unendliche Menge in L enthalten ist. Diese Eigenschaft gilt aber nur für hinreichend lange Worte. Für kürzere Worte ist nichts ausgesagt.

   Für unsere Sprache L könnte nun $n > 2$ sein. In diesem Fall gäbe es kein Wort z, für das etwas zu zeigen wäre, denn es gibt ja kein Wort z mit $|z| \geq n > 2$, und das PL gilt trivialerweise.

Übungsaufgabe 2.3: Gegeben ist der folgende ϵ-FA A. Berechnen Sie für A die ϵ-Hülle und konstruieren Sie mit dem Verfahren der Vorlesung den zu A äquivalenten ϵ-freien NFA.

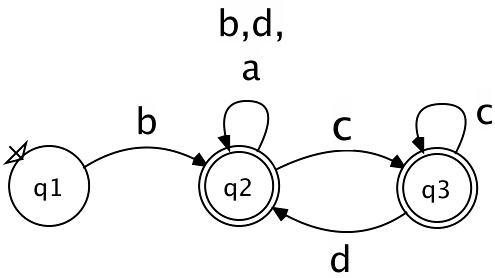
von
2



Lösung:

$$R_A = \{(q_2, q_1), (q_2, q_3)\} \cup id_Q$$

Die hinzugefügten Kanten sind durch etwas größere Anschriften dargestellt. Man beachte, dass q_2 zum Endzustand wird.



von
4

Übungsaufgabe 2.4:

1. Sei $w \in \{0,1\}^*$, dann bezeichnet $\bar{w}$ das Wort, das man erhält, indem man in w jede 0 durch 1 ersetzt (und umgekehrt). Bsp. $\overline{100} = 011$.

   Zeigen Sie mit Hilfe des Pumping-Lemmas, dass die Sprache $L = \{w\bar{w} \mid w \in \{0,1\}^*\}$ nicht regulär ist.

   Lösung: Eingangsfeststellung: Jedes Wort $w\bar{w}$ aus L besitzt genausoviele 0 wie 1, denn für jedes Symbol in w steht in $\bar{w}$ das jeweils andere.

   Angenommen die Sprache L wäre regulär, dann wäre das Pumping-Lemma anwendbar.

   Pumping Lemma: Sei L eine reguläre Sprache. Dann existiert eine Zahl n, so dass für alle $z \in L$ mit $|z| \geq n$ stets eine Zerlegung $z = uvw$ existiert, so dass gilt:
   (i) $|uv| \leq n$
   (ii) $|v| \geq 1$
   (iii) $\forall i \in \mathbb{N} : uv^iw \in L$

   Sei n die Zahl des PL. Betrachte das hinreichend lange Wort $z = 0^n1^n$. Dieses ist in L, da $\overline{0^n} = 1^n$ ist. Es muss also eine Zerlegung $z = uvw$ mit obigen Eigenschaften existieren.

   Da die ersten n Symbole von z nur aus 0 bestehen, folgt aus (i), dass u, v aus $\{0\}^*$ sind, d.h. es gibt ein l, so dass $u = 0^l$, und ein k, so dass $v = 0^k$, und wegen (ii) gilt $k > 0$.

   Wir betrachten nun das Wort $z' = uv^iw$ für $i = 0$, das nach PL auch in L sein muss:

   $$z' = uv^0w = 0^l0^{ik}0^{n-k-l}1^n = 0^{n-k}1^n$$

   Da $k > 0$, besitzt $z' = 0^{n-k}1^n$ nicht genausoviele 0 wie 1, d.h. z' kann aufgrund der Eingangsbemerkung nicht in L sein

   Damit ist z' nicht in L. Widerspruch zur Annahme, dass L regulär ist.

von
6

Übungsaufgabe 2.5:

1. Sei $L \subseteq \Sigma^*$ eine beliebige Sprache und $a \in \Sigma$. Definiere:

$$(L\%a) := \{w \in \Sigma^* \mid \text{ es gibt ein Wort } wa \text{ in } L\ \}$$

$L\%a$ entsteht also aus L, wenn man nur auf a endende Worte aus L betrachtet und bei denen dieses letzte a streicht.

Zeige: Wenn $L \subseteq \Sigma^*$ eine beliebige reguläre Sprache ist, dann ist auch $(L\%a)$ regulär.

Lösung: Sei A der DFA, der L akzeptiert. (Diesen DFA muss es stets geben, da L regulär ist.)

Wir konstruieren einen neuen NFA B, der zunächst die gleichen Zustände und Kanten wie A hat. Auch der Startzustand ist gleich.

Wir fügen einen neuen Zustand q_{neu} hinzu. Dieser Zustand ist der einzige Endzustand von B.

Wir betrachten nun alle Kantenzüge in A, die die Länge 2 haben, deren zweite Kante mit einem a beschriftet ist und deren letzter Zustand ein Endzustand ist ($p, q, r \in Q, x \in \Sigma$):

$$p \xrightarrow{x} q \xrightarrow{a} r \in F$$

In diesem Fall fügen wir in B die Kante $p \xrightarrow{x} q_{neu}$ auf den neuen Endzustand hinzu. (Dieser Schritt macht B i.a. zu einem NFA.)

- Der Automat B hat nun die Möglichkeit, wann immer A mit einem Wort der Form wa in einen Endzustand gelangt, d.h. wenn gilt

  $$q_0 \xrightarrow{w_1} q_1 \cdots \xrightarrow{w_n} q_n \xrightarrow{a} q_{n+1},$$

  ebenfalls zu akzeptieren:

  $$q_0 \xrightarrow{w_1} q_1 \cdots \xrightarrow{w_{n-1}} q_{n-1} \xrightarrow{w_n} q_{neu},$$

  denn in B haben wir für diesen Fall ja die Kante $q_{n-1} \xrightarrow{w_n}\in q_{neu}$ hinzugefügt. Also gilt $wa \in L(A) \Longrightarrow w \in L(B)$
- Umgekehrt sei nun $w \in L(B)$. Da wir nur den einen Endzustand haben, muss es einen Pfad geben, der in q_{neu} endet:

  $$q_0 \xrightarrow{w_1} q_1 \cdots \xrightarrow{w_{n-1}} q_{n-1} \xrightarrow{w_n} q_{neu},$$

  Da wir in B die Kante $q_{n-1} \xrightarrow{w_n} q_{neu}$ nur dann hinzugefügt haben, wenn wir in A einen Kantenzug der Form $q_{n-1} \xrightarrow{x} q \xrightarrow{a} r$ mit $q \in Q$ und $r \in F$ vorfinden, wissen wir, dass wir in A mit wa den Pfad

  $$q_0 \xrightarrow{w_1} q_1 \cdots \xrightarrow{w_{n-1}} q_{n-1} \xrightarrow{w_n} q \xrightarrow{a} r$$

  durchlaufen würden, und dies führt in A zur Akzeptierung. Also $w \in L(B) \Longrightarrow wa \in L(A)$.

Also gilt $wa \in L(A) \iff w \in L(B)$, d.h. $L(B) = L\%a$. Also ist $L\%a$ auch regulär.

Da wir zu jeder regulären Sprache L den Automaten B konstruieren können, wissen wir dass $L\%a$ für jede reguläre Sprache L auch regulär ist.

2. Sei $L \subseteq \Sigma^*$ eine reguläre Sprache. Zeigen Sie, dass dann auch die Menge der *kürzesten Worte (KW)*:

$$KW(L) \quad := \quad \{w \in L \mid \text{ kein echtes Anfangsstück von } w \text{ ist auch in } L\ \}$$

eine reguläre Sprache ist.

Lösung: Da L regulär ist, existiert ein vollständiger DFA A, der L akzeptiert. Sei $A = (Q, \Sigma, \delta_A, z_0, F)$.

Wir definieren den NFA $B = (Q, \Sigma, \delta_B, \{z_0\}, F_B)$, indem wir in A jede Kante in einen Endzustand hinein auf einen neuen Zustand f „umbiegen", der dann der einzige Endzustand ist. Genauer: Wir entfernen die alte Kante und fügem die neue hinzu. WIr setzen f als einizigen Endzustad in B.

Von diesem Endzustand kommt man mit jedem Symbol zu einem weiteren Zustand q_{neu}, und von q_{neu} kommt man mit jedem Symbol wieder zu q_{neu}.

Wir zeigen $L(B) = KW(L)$.

(a) $KW(L) \subseteq L(B)$:
Betrachten wir ein Wort $w \in KW(L)$. Der Automat A absolviert für w folgende Zustandsfolge:

$$q_0 \xrightarrow{x_1} q_1 \cdots \xrightarrow{x_n} q_n \in F_A$$

und da kein echtes Anfangsstück von w auch in L ist und A ein DFA ist, ist keiner der Zuständ $q_0, \ldots, q_{n-1}$ ein Endzustand in A. Daraus folgt, dass der Automat B fast die gleiche Zustandsfolge durchläuft, nur dass der letzte Zustand (nach Konstruktion) jetzt f ist. Insbesondere wird w aber auch in B akzeptiert.

(b) $L(B) \subseteq KW(L)$.
Betrachten wir ein akzeptiertes Wort $w = x_1 \cdots x_n$ mit $x_i \in \Sigma$. Der Automat B absolviert folgende Zustandsfolge:

$$q_0 \xrightarrow{x_1} q_1 \cdots \xrightarrow{x_n} q_n$$

Nach Konstruktion von δ_B ist $q_n = f$, denn $w \in L(B)$. Die Zuständ $q_0, \ldots, q_{n-1}$ müssen dann aus Q_A sein, wären also so auch in A möglich gewesen. Wäre nun Anfangsstück von w in A akzeptiert worden (bspw. $x_1 \cdots x_k$ mit $k < n$), dann wäre bereits $q_k = f$ und $q_{k+1} = q_{k+2} = \cdots = q_n = q_{neu}$ und B würde w gar nicht akzeptieren. Also gibt es kein echtes Anfangsstück von w, dass von A akzeptiert wird.