FGI-2 – Formale Grundlagen der Informatik II

Modellierung und Analyse von Informatiksystemen

Musterlösung 1 : Endliche Automaten

Präsenzteil am 14./15. 10. – Abgabe am 21./22. 10. 2013

Präsenzaufgabe 1.1:

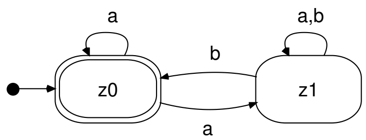
1. Wir wissen aus FGI-1, dass es zu jedem NFA A einen DFA B mit $L(A) = L(B)$ gibt. Kann man B aus A berechnen? Wenn ja, wie?

   Lösung: Potenzautomatenkonstruktion (siehe FGI-1).

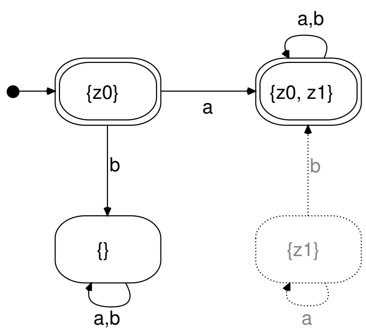
2. Sei A ein NFA mit $L(A) \subseteq X^*$. Geben Sie eine Konstruktionsvorschrift für einen NFA $\bar{A}$ an, für den $L(\bar{A}) = X^* \setminus L(A)$ gilt. (Tipp: Wandeln Sie zunächst A in einen DFA um.)

   Lösung: O.b.d.A. sei A ein vollständiger DFA. Wir konstruieren $\bar{A}$ aus A, indem wir in $\bar{A}$ genau die Zustände als Endzustände wählen, die dies in A nicht sind. Die restlichen Komponenten bleiben gleich. Auch $\bar{A}$ ist ein vDFA. Wurde in A ein Wort akzeptiert, dann in $\bar{A}$ nicht und umgekehrt.

3. Konstruieren Sie den Potenzautomaten für folgenden NFA:

   Lösung: Der Potenzautomat ergibt sich wie folgt (der gestrichelte Teil gehört nicht zur initialen Zusammenhangskomponente und braucht nicht erzeugt zu werden):

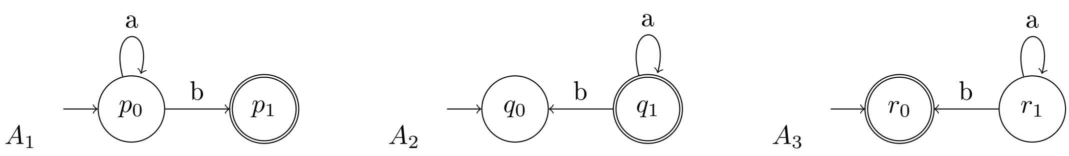
Präsenzaufgabe 1.2: Zeigen Sie die erste Teilaussage von Lemma 1.9: „Das Leerheitsproblem für NFA ist entscheidbar.“

1. Beschreiben Sie ein Verfahren, welches für einen gegebenen nicht-deterministischen Automaten $A := (Q, \Sigma, \delta, Q^0, F)$ feststellt, ob $L(A) = \emptyset$ gilt.

Lösung: Zuerst ist die initiale Zusammenhangskomponente von A zu bilden. Hierzu wird die Menge der erreichbaren Zustände Q' zunächst mit den Startzuständen von A gefüllt ($Q' := Q^0$) und anschließend so lange um Nachfolgezustände von allen bereits in Q' enthaltenen Zuständen ergänzt, bis keine Änderungen mehr auftreten.

Enthält Q' keinen Endzustand (es gilt $Q' \cap F = \emptyset$), so akzeptiert A kein einziges Wort, also gilt $L(A) = \emptyset$.

2. Wenden Sie Ihr Verfahren auf folgende Automaten an:



Lösung: $p_1 \in Q_1' = \{p_0, p_1\}$ $\quad$ $q_1 \notin Q_2' = \{q_0\}$ $\quad$ $r_0 \in Q_3' = \{r_0\}$

$L(A_1) = \{a\}^*\{b\} \neq \emptyset$ $\quad$ $L(A_2) = \emptyset$ $\quad$ $L(A_3) = \{\varepsilon\} \neq \emptyset$

3. Begründen Sie Korrektheit und Termination Ihres Verfahrens.

Lösung: *Korrektheit:* Die Argumentation wird leichter, wenn beide Seiten negiert werden:

$$\begin{array}{rlcl} & L(A) = \emptyset & \Leftrightarrow & Q' \cap F = \emptyset \\ \equiv & L(A) \neq \emptyset & \Leftrightarrow & Q' \cap F \neq \emptyset \end{array}$$

Es sind nun zwei Richtungen zu zeigen:

(a) Falls $L(A) \neq \emptyset$, dann $Q' \cap F \neq \emptyset$.
Sei $w = a_1 a_2 \ldots a_n \in L(A)$. Dann gibt es eine Rechnung von einem Startzustand $q_0 \in Q^0$ zu einem Endzustand $q_e \in F$: $q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \ldots \xrightarrow{a_n} q_e$. Da q_{i+1} jeweils ein Nachfolgezustand von q_i ist und q_0 anfänglich in Q' aufgenommen wurde, ist am Ende auch $q_e \in Q'$. Also ist die Schnittmenge $Q' \cap F$ nicht leer.

(b) Falls $Q' \cap F \neq \emptyset$, dann $L(A) \neq \emptyset$.
Sei $q_e \in Q' \cap F$. Dann ist q_e entweder ein Startzustand (direkt in Q' enthalten) oder ein Nachfolgezustand eines zuvor schon in Q' aufgenommenen Zustandes q'. Für q' gilt Gleiches: Entweder ist $q' \in Q^0$ oder q' ist Nachfolgezustand von q''. Diese Argumentation lässt sich fortsetzen, bis man auf einen Startzustand q^m trifft. Die Übergänge $q^m \xrightarrow{a_1} \ldots \xrightarrow{a_{m-1}} q' \xrightarrow{a_m} q_e$ bilden eine Erfolgsrechnung im NFA A, so dass $w = a_1 a_2 \ldots a_n \in L(A)$ gilt. Also ist $L(A)$ nicht leer.

Termination: Die Konstruktion der Menge Q' bricht garantiert nach endlich vielen Schritten ab, da die Menge Q, aus welcher die Elemente für Q' stammen, endlich ist. Ob Q' einen Endzustand enthält, lässt sich durch reihenweises Überprüfen der in Q' enthaltenen Elemente in endlicher Zeit feststellen. Somit terminiert das Verfahren immer in endlicher Zeit.

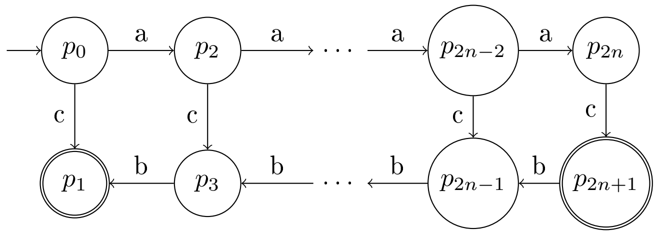
4. Ist Ihr Verfahren ohne Modifikationen für deterministische und verallgemeinerte endliche Automaten anwendbar? Wenn nicht, was müsste modifiziert werden?

Lösung: Das Verfahren für NFA ist genauso für DFA anwendbar, da ein DFA lediglich ein Spezialfall eines NFA ist.

Die Argumentation zu Teilaufgabe 3 ist nicht davon abhängig, ob an den Kanten genau ein Symbol steht. Beim verallgemeinerten FA können die Teile $a_1 \ldots a_n$ als Teilwörter beliebiger Länge betrachtet werden.

von 6

Übungsaufgabe 1.3: Gegeben ein beliebiges, festes n. Zu n sei der Automat A_n wie folgt gegeben:



1. Beschreiben Sie $L(A_n)$ durch einen regulären Ausdruck. (Sie dürfen Auslassungspunkte verwenden.)

   Lösung: $c + (acb) + (a^2cb^2) + \cdots + (a^ncb^n) + (a^nc)$

2. Beschreiben Sie $L(A_n)$ als Menge (ohne Auslassungspunkte).

   Lösung: $L(A_n) = \{a^icb^i \mid 0 \leq i \leq n\} \cup \{a^nc\}$

3. Begründen Sie die Korrektheit Ihrer Antworten.

   Lösung: Die Endzustände liegen in der unteren Reihe. Um vom Startzustand (in der oberen Reihe) zur unteren Reihe zu kommen, muss einmalig c gelesen werden. Abhängig davon, von welchem Zustand p_{2i} aus der c-Übergang gewählt wurde, sind vorher i a-s und hinterher auch i b-s zu lesen. Somit akzeptiert A_n genau jene Wörter, die aus einer Anzahl von a-s vorne und genau so vielen b-s hinten bestehen, in der Mitte getrennt durch ein c. Hinzu kommt die Möglichkeit, bei n a-s nach dem c im Zustand p_{2n+1} zu halten.

   Im regulären Ausdruck bleibt keine andere Möglichkeit, als alle diese Wörter einzeln als Alternativen aufzulisten. Die Mengenschreibweise bietet eine kompaktere Form.

4. Ist $L(A_n)$ regulär? Begründen Sie dies.

   Lösung: Natürlich ist $L(A_n)$ regulär, da die Sprache von einem endlichen Automaten akzeptiert wird und durch einen regulären Ausdruck beschrieben werden kann. Außerdem ist jede endliche Menge (wie $L(A_n)$) regulär.

5. Zusatz: Betrachten Sie die Vereinigung aller Mengen $L(A_n)$: $A := \bigcup_{n\geq 0} L(A_n)$

   Ist A regulär? Kann man A durch eine Grammatik darstellen? Begründen Sie dies.

   Lösung: $A = \{a^icb^i \mid 0 \leq i\} \cup \{a^ic \mid 0 \leq i\}$ ist eine nicht-reguläre, aber kontextfreie Menge und kann durch eine kontextfreie Grammatik mit folgenden Produktionsregeln dargestellt werden: $P = \{S \to T \mid U, T \to aTb \mid c, U \to aU \mid c\}$. Es kann erwähnt werden, dass man das mit dem „Pumping Lemma“ zeigen kann (als Wort a^mcb^m wählen, wobei m die Zahl aus dem Pumping Lemma ist. Dann ist bei jeder Zerlegung in uvw das v komplett in den a's enthalten und wenn man v „wegfallen“ läßt, ist das Wort nicht mehr in A).

Übungsaufgabe 1.4: Es soll ein einfacher Textverarbeitungs-Algorithmus in Form eines endlichen Automaten erstellt werden. Gesucht ist ein endlicher Automat, der die akzeptierten Wörter eines anderen Automaten akzeptiert, mit dem Unterschied das jedes Symbol verdoppelt wird.

von
6

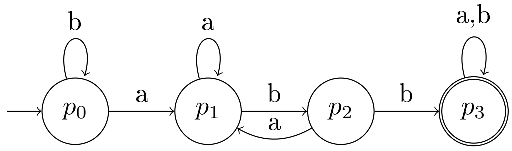
1. Beschreiben Sie ein Verfahren, welches für einen gegebenen deterministischen Automaten $A := (Q, \Sigma, \delta, \{q^0\}, F)$ einen deterministischen Automaten $A' := (Q', \Sigma', \delta', \{q'^0\}, F')$ konstruiert, der genau die 'gedoppelten' Wörter von $L(A)$ akzeptiert,
   d.h. $L(A') = \{a_0a_0a_1a_1 \dots a_na_n | a_0a_1 \dots a_n \in L(A)\}$.

   Lösung: Folgende Schritte liefern das gewünschte Ergebnis:

   (a) In DFA A die Übergänge aufteilen, also für jeden Übergang einen eigenen Pfeil zeichnen.

   (b) Jetzt jeden Übergang / Pfeil durch einen neuen Zustand ersetzten, von dem ein gleicher Übergang / Pfeil abgeht und auch ankommt. Dabei ist darauf zu achten, dass die neuen Zustände nie Endzustände sind.

2. Beweisen Sie, dass der folgende Automat A, die Menge aller Wörter über $\Sigma = \{a, b\}$ akzeptiert, die das Teilwort abb enthalten.



A

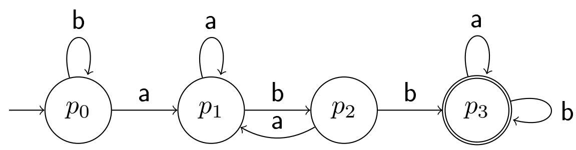
   Lösung: Alle Wege von p_0, die zum ersten Mal auf p_3 treffen, haben zuletzt 3 Kanten mit dem Wort abb durchlaufen. In p_3 können dann allerdings noch beliebige Worte aus a und b folgen, und da von allen Zuständen aus sowohl a als auch b Eingaben verarbeitet werden können, und man von jedem Zustand aus mit abb nach p_3 kommt, tut der Automat genau das geforderte.

3. Beschreiben Sie diese Wortmenge durch einen regulären Ausdruck.

   Lösung: $L(A) = (a + b)^*abb(a + b)^*$

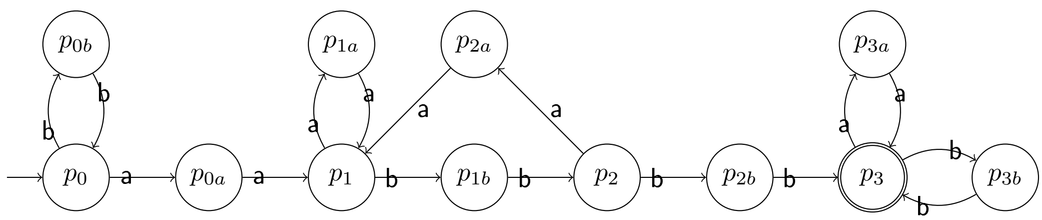
4. Wenden Sie Ihr Verfahren unter 1. auf diesen Automaten A an. Dokumentieren Sie dabei die Zwischenergebnisse.

   Lösung: Nach Schritt (a) ergibt sich folgender DFA B:



B

   Mit Schritt (b) erhalten wir den DFA A':



A'

5. Begründen Sie die Korrektheit der Lösung.

Lösung: Alle Wege vom Anfangszustand p_0, die zum ersten Mal auf den Endzustand p_3 treffen, haben zuletzt 6 Kanten mit dem Wort $aabbbb$ durchlaufen, da man von p_2 nach p_3 nur bb durchlaufen kann, und von p_1 nach p_2 ebenfalls nur bb durchlaufen kann, und nach p_1 nur mit dem Durchlaufen von aa kommt. Das $aabbbb(aa+bb)^* \subset L(A')$ ist wegen p_{3a} und p_{3b} klar, das tatsächlich $L(A') = (aa+bb)^* aabbbb(aa+bb)^*$ gilt, überlegt man sich leicht, da man von p_0, p_1 und p_2 aus je aa oder bb durchlaufen kann und wieder bei einem dieser Zustände landet, wenn man nicht auf p_3 stößt.

Bisher erreichbare Punktzahl: 12