

# Grundlagen der theoretischen Informatik, SS 2001

## 1. Teilklausur, Montag, 11. Juni 2001 — Lösungen

*Anleitung:* Bearbeiten Sie jede Aufgabe auf dem dafür vorgesehenen *Blatt*. Verwenden Sie gegebenenfalls Ergänzungsblätter. Schreiben Sie auf *alle* Blätter Ihren Namen und Ihre Matrikelnummer.

Sie können alle Aussagen, die in der Vorlesung oder auf den Übungsblättern vorgekommen sind, verwenden, ohne die Beweise abzuschreiben.

Kriterium für positive Bewertung: mindestens 50%, das sind 15 Punkte.

1. (10 Punkte) Zeichnen Sie das Zustandsdiagramm des nichtdeterministischen endlichen Automaten  $(\{a, b, c, d\}, \{0, 1\}, \delta, \{a\}, \{d\})$  mit

$$\delta = \{(a, 0, b), (a, 1, a), (a, 1, c), (b, 0, a), (b, 1, b), (b, 1, c), (c, 0, d), (d, 0, d)\},$$

und bestimmen Sie den Minimalautomaten.

**Lösung:** Nebenstehend ist der deterministische Automat angegeben, der aus der Potenzmengenkonstruktion entsteht, wobei mit den Einzelzuständen begonnen wurde, aber nur diejenigen zusätzlichen Zustände berücksichtigt wurden, die auf der rechten Seite auftreten.

$\delta'$	0	1
a	b	ac
b	a	bc
c	d	$\emptyset$
d	d	$\emptyset$
ac	bd	ac
bc	ad	bc
ad	bd	ac
bd	ad	bc
$\emptyset$	$\emptyset$	$\emptyset$

Die Zustände  $c, d, \emptyset$  sind nicht erreichbar und können gestrichen werden. Es bleibt die Zustandsmenge  $Q = \{a, b, ac, bc, ad, bd\}$  mit den akzeptierenden Zuständen  $F = \{ad, bd\}$ .

Anfangszerlegung der Zustandsmenge:  $K_1 = Q - F = \{a, b, ac, bc\}$ ,  $K_2 = F = \{ad, bd\}$ .  $K_1$  wird zerlegt gemäß  $\delta(a, 0), \delta(b, 0) \in K_1$ ,  $\delta(ac, 0), \delta(bc, 0) \in K_2$ .

Neue Zerlegung  $K_1 = \{a, b\}$ ,  $K_2 = \{ac, bc\}$ ,  $K_3 = \{ad, bd\}$ . Keine weitere Zerlegung möglich. Der Minimalautomat ist  $(\{q_1, q_2, q_3\}, \{0, 1\}, \bar{\delta}, q_1, \{q_3\})$  mit  $\bar{\delta} = \{(q_1, 0, q_1), (q_1, 1, q_2), (q_2, 0, q_3), (q_2, 1, q_2), (q_3, 0, q_3), (q_3, 1, q_2)\}$ .

2. (10 Punkte)

- (a) (7 Punkte) Konstruieren Sie einen endlichen Automaten für die Sprache  $L_1$  der Wörter über  $\{0, 1\}$ , bei denen auf jedes Vorkommen der Teilfolge 0011 unmittelbar die Teilfolge 1100 anschließt.

Zum Beispiel ist  $w_1 = 0001111100 \notin L_1$ ,  $w_2 = 01011001111001111000 \in L_1$ ,  $w_3 = 0010 \in L_1$ , und  $w_4 = 00111100011 \notin L_1$ . Erklären Sie Ihren Automaten in wenigen Sätzen. (Ein formaler Korrektheitsbeweis ist nicht erforderlich.)

Geben Sie eine akzeptierende Zustandsfolge für  $w_2$  und  $w_3$  an.

- (b) (3 Punkte) Konstruieren Sie einen endlichen Automaten, der alle *Präfixe* der Wörter aus  $L_1$  akzeptiert.

$$L_2 = \{x \mid \text{es gibt ein } y \text{ mit } xy \in L_1\}$$

(Zum Beispiel ist  $w_5 = 010110011110011110 \in L_2 - L_1$ , wegen  $w_5 00 = w_2 \in L_1$ .)

**Lösung zu a:** Wir konstruieren einen DEA. (Einen NEA zu finden, der den Nicht-determinismus ausnützt, ist schwierig.) Der Automat merkt sich den längsten Präfix der Wortes 0011, mit dem die bisher gelesene Eingabe aufhört. Wenn die letzten  $i$  gelesenen Symbole mit den ersten  $i$  Symbolen von 0011 übereinstimmen, befindet sich der Automat im Zustand  $q_i$ .

Im Zustand  $q_4$  hat der Automat erkannt, dass die Folge 0011 vorgekommen ist, und er prüft nun, ob die Symbole 1100 in dieser Reihenfolge kommen, und geht dabei in die Zustände  $q_5, q_6, q_7$ . Wenn das Ende der Folge 00111100 erreicht ist, geht der Automat wieder in den Zustand  $q_2$ . Wenn die Folge 1100 nicht ohne Fehler gefunden wird, geht der Automat in einen Abbruchzustand  $q_\infty$ , in dem er immer bleibt.

$\delta$	0	1
$q_0$	$q_1$	$q_0$
$q_1$	$q_2$	$q_0$
$q_2$	$q_2$	$q_3$
$q_3$	$q_1$	$q_4$
$q_4$	$q_\infty$	$q_5$
$q_5$	$q_\infty$	$q_6$
$q_6$	$q_7$	$q_\infty$
$q_7$	$q_2$	$q_\infty$
$q_\infty$	$q_\infty$	$q_\infty$

Der Ausgangszustand ist  $q_0$ , die akzeptierenden Zustände sind  $F = \{q_0, q_1, q_2, q_3\}$ .

Der Weg von  $q_0$  bis  $q_7$  (oder gegebenenfalls  $q_8$ ) ist klar, bei der „Rückwärtskanten“, wenn nicht das erwartete Zeichen kommt, muss man etwas aufpassen, wohin sie führen. (Zum Beispiel ist  $\delta(q_2, 0) = q_2$  weil im Zustand  $q_2$  nach dem Lesen einer 0 bekannt ist, dass das Wort mit 000 aufhört; Die letzten beiden Nullen sind der längste Präfix von 0011 am Ende des bisher gelesenen Wortes.)

Zur Kontrolle: Bei Eingabe von  $w_2$  durchläuft der endliche Automat die akzeptierende Zustandsfolge  $q_0 q_1 q_0 q_1 q_0 q_0 q_1 q_2 q_3 q_4 q_5 q_6 q_7 q_2 q_3 q_4 q_5 q_6 q_7 q_2 q_2$ , bei  $w_3$  die Zustandsfolge  $q_0 q_1 q_2 q_3 q_1$ .

**Lösung zu b:** Man muss die Endzustände des vorigen endlichen Automaten auf  $F' = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7\}$  erweitern. Wenn der Automat in einem der neuen Endzustände  $q_4, q_5, q_6, q_7$  ist, kann er durch Verlängerung des Eingabewortes (um einen passenden Suffix von 1100) in den für den ursprünglichen Automaten akzeptierenden Zustand  $q_2$  gebracht werden. Daher ist das gelesene Wort ein Präfix eines Wortes aus  $L_1$ .

Wenn der Automat im Zustand  $q_\infty$  ist, dann hilft auch eine Verlängerung des gelesenen Wortes nichts mehr, um in einen akzeptierenden Zustand zu kommen. Das gelesene Wort ist kein Präfix eines Wortes aus  $L_1$ .

3. (10 Punkte) Beweisen Sie:

- (a) (5 Punkte) Wenn  $L_1$  regulär ist, dann ist auch die Sprache  $L_3$  aller *Suffixe* von  $L_1$  regulär.

$$L_3 = \{x \mid \text{es gibt ein } y \text{ mit } yx \in L_1\}$$

- (b) (5 Punkte) Wenn  $L_1$  regulär ist, dann ist auch die Sprache  $L_4$  aller Teilwörter von  $L_1$  regulär.

$$L_4 = \{x \mid \text{es gibt } y, z \text{ mit } yxz \in L_1\}$$

**Lösung 1 zu a:** Die Lösung von Aufgabe 2b lässt sich verallgemeinern. Wenn  $L(M) = L_1$  ist, erhält man einen Automaten  $M'$  für die Sprache  $L_2$  der *Präfixe* von  $L_1$  indem man in  $M$  alle Zustände, von denen ein akzeptierender Zustand von  $M$  erreichbar ist, zusätzlich zu akzeptierenden Zuständen macht.

Beweis für den Fall eines deterministischen Automaten  $M$ : (Das Verfahren funktioniert aber genauso gut bei nichtdeterministischen endlichen Automaten.) Wenn  $\delta(q_0, x) = q$  ist und von  $q$  aus ein akzeptierender Zustand  $q_F \in F$  erreichbar ist ( $\delta(q, y) = q_F$  für ein  $y \in \Sigma^*$ ), dann ist  $\delta(q_0, xy) = q_F \in F$  und somit  $xy \in L_1$ , und  $x \in L_2$ . Umgekehrt, wenn es ein  $y \in \Sigma^*$  mit  $xy \in L_1$  gibt, dann ist  $\delta(q_0, xy) = \delta(\delta(q_0, x), y) \in F$  und somit ist von  $\delta(q_0, x)$  aus ein akzeptierender Zustand erreichbar. Somit ist  $\delta(q_0, x) \in F'$  und  $x$  wird von  $M'$  akzeptiert.

Reguläre Sprachen sind abgeschlossen unter *Umkehrung* (oder Spiegelbild, siehe Übungsaufgabe 25). Die Suffixe eines Wortes oder einer Sprache sind die Umkehrungen der Präfixe der Umkehrung:

$$L_3 = \text{Suffixe}(L_1) = (\text{Präfixe}(L_1^R))^R$$

Somit ergibt sich die Abgeschlossenheit gegenüber Suffixen aus der Abgeschlossenheit gegenüber Präfixen.

**Lösung 2:** Man kann auch direkt einen nichtdeterministischen endlichen Automaten für die Suffixe konstruieren. Man macht alle Zustände, die von einem Ausgangszustand aus erreichbar sind, zusätzlich zu Ausgangszuständen. Der Beweis ist analog wie bei Lösung 1.

**Lösung 3:** Eine andere Möglichkeit ist ein nichtdeterministischer endlicher Automat mit  $\varepsilon$ -Übergängen:  $M$  sei ein deterministischer oder nichtdeterministischer Automat mit  $L(M) = L_1$ . Man erstellt eine Kopie  $M'$  von  $M$  und ersetzt alle Übergänge  $(q'_1, a, q'_2) \in \delta'$  durch  $\varepsilon$ -Übergänge  $(q'_1, \varepsilon, q'_2)$ . Zusätzlich verbindet man jeden Zustand  $q' \in Q'$  des Automaten  $M'$  mit dem entsprechenden Zustand  $q \in Q$  durch einen  $\varepsilon$ -Übergang  $(q', \varepsilon, q)$ . Der neue Automat  $M_3$ , der  $L_3$  akzeptiert, hat die Zustände  $Q \cup Q'$ , die Anfangszustände  $Q'_0 \subseteq Q'$  und die akzeptierenden Zustände  $F$  des ursprünglichen Automaten. Dieser Automat kann nun mit den  $\varepsilon$ -Übergängen in  $Q'$  beginnen und so tun, als ob er ein Wort  $y$  liest, ohne es wirklich zu lesen. Irgendwann entscheidet er sich, in die Zustandsmenge  $Q$  zu springen und den Rest (Suffix)  $x$  des Wortes  $yx \in L_1$  wirklich zu lesen.

**Lösung zu b:** Die Teilwörter sind die Präfixe der Suffixe (oder die Suffixe der Präfixe). Somit ergibt sich der Beweis aus Lösung 1 zu Teil (a).

4. (Zusatzfrage, 5 Punkte) Wie kann man für zwei gegebene reguläre Sprachen  $L_1$  und  $L_2$  entscheiden, ob  $L_1 \subseteq L_2$  ist?

**Lösung:**  $L_1 \subseteq L_2$  gilt genau dann, wenn  $L_1 \cap \bar{L}_2 = \emptyset$  ist. Reguläre Sprachen sind gegenüber Durchschnitt und Komplement abgeschlossen. Daher braucht man nur zu überprüfen, ob die reguläre Sprache  $L_1 \cap \bar{L}_2$  leer ist.