

---

## Einführung in die Theoretische Informatik

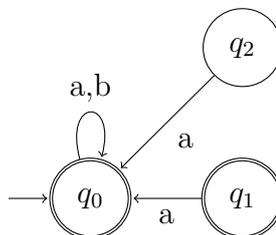
---

### Hausaufgabe 1 (5 Punkte)

1. Konstruieren Sie ein Beispiel dafür, dass unerreichbare Zustände nicht automatisch äquivalent sind. Begründung!
2. Sei  $\Sigma$  ein Alphabet und  $L \subseteq \Sigma^*$ . Beweisen oder widerlegen Sie:
  - (a) Wenn  $w \equiv_L w'$  und  $v \in \Sigma^*$ , dann ist auch  $wv \equiv_L w'v$ .
  - (b) Wenn  $w \equiv_L w'$  und  $v \in \Sigma^*$ , dann ist auch  $vw \equiv_L vw'$ .

### Lösungsvorschlag

1. Es genügt ein jeweils unerreichbarer Endzustand und ein Nicht-Endzustand. Z.B. gilt  $q_1 \not\equiv_A q_2$  in folgendem DFA  $A$ :



2. (a) Wir nehmen an, dass  $w \equiv_L w'$ , und damit  $\forall z.(wz \in L \leftrightarrow w'z \in L)$ .  
Für  $wv \equiv_L w'v$  sei  $u \in \Sigma^*$  beliebig. Dann folgt  $wvu \in L \leftrightarrow w'vu \in L$  direkt aus der Annahme, wenn man  $vu$  für  $z$  einsetzt.
- (b) Die Aussage ist falsch. Gegenbeispiel:  $L = L(a^*b^*)$ ,  $w = ab$ ,  $w' = b$ ,  $v = b$ .

### Hausaufgabe 2 (5 Punkte)

Beschreiben Sie ein Entscheidungsverfahren, welches für einen gegebenen DFA  $M = (Q, \Sigma, \delta, q_0, F)$  entscheidet, ob jedes von  $M$  akzeptierte Wort eine ungerade Länge hat. Benutzen Sie dabei den Test auf Leerheit der Sprache eines Automaten.

## Lösungsvorschlag

Man konstruiert einen DFA  $G$  für die Sprache  $(\Sigma\Sigma)^* = \{w \in \Sigma^* \mid |w| \text{ ist gerade}\}$ . Dann konstruiert man einen Produktautomat  $P$ , der die Sprache

$$L(M) \cap L(G) = \{w \in L(M) \mid |w| \text{ ist gerade}\}$$

akzeptiert (vgl. Satz 2.20). Der ursprüngliche Automat hat die gefragte Eigenschaft genau dann wenn  $L(P) = \emptyset$ , was man mit dem Leerheitstest feststellen kann.

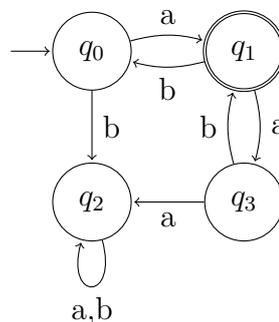
Ein anderes Verfahren arbeitet direkt auf dem ursprünglichen Automaten: Man markiert iterativ alle Zustände, die mit einem Eingabewort gerader Länge erreicht werden können. Dafür beginnt man mit dem Startzustand und markiert dann jeweils ausgehend von einem bereits markierten Zustand  $q$  die Zustände  $\{\hat{\delta}(q, ab) \mid a, b \in \Sigma\}$ , bis ein Fixpunkt erreicht ist. (Dies ist eine einfache Variante der Tiefensuche.) Der Automat hat die gefragte Eigenschaft genau dann wenn am Schluss kein Endzustand markiert ist.

## Hausaufgabe 3 (5 Punkte)

Sei  $\Sigma$  ein Alphabet. Für eine Sprache  $L \subseteq \Sigma^*$  definieren wir die folgenden Sprachen:

$$\text{suff}(L) = \{v \in \Sigma^* \mid \exists w \in \Sigma^*. vw \in L\}$$

1. Geben Sie einen regulären Ausdruck für  $\text{suff}(L((ab)^*b))$  an.
2. Gegeben sei folgender DFA  $M$ :



Geben Sie einen endlichen Automaten (DFA, NFA oder  $\epsilon$ -NFA) an, der  $\text{suff}(L(M))$  akzeptiert.

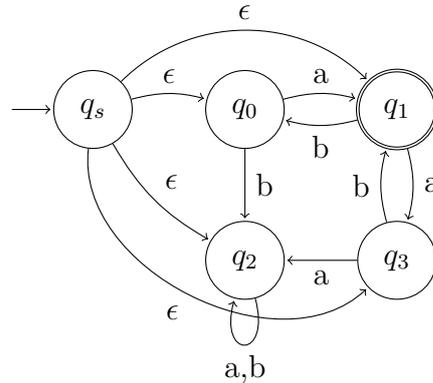
3. Zeigen Sie: Wenn  $L \subseteq \Sigma^*$  regulär ist, so ist auch  $\text{suff}(L)$  regulär.

*Hinweis, falls Sie eine Automatenkonstruktion verwenden:* Beschreiben Sie Ihre Konstruktionsidee zunächst informell und geben Sie danach den Automaten formal als 5-Tupel an.

## Lösungsvorschlag

1.  $\epsilon \mid (b \mid \epsilon)(ab)^*b$

2.



3. Idee: Sei  $N$  ein  $\epsilon$ -NFA  $N = (Q, \Sigma, \delta, q_0, F)$ , bei dem jeder Zustand vom Startzustand erreichbar ist. Daraus konstruiert man einen neuen  $\epsilon$ -NFA  $N'$  mit  $L(N') = \text{suffix}(L(N))$ . Dafür fügt man einen neuen Startzustand hinzu, von dem ein  $\epsilon$ -Übergang in jeden Zustand führt.

Formal ist  $N'$  wie folgt definiert:  $N' = (Q \cup \{q_s\}, \Sigma, \delta', q_s, F)$ , mit

$$\begin{aligned} \delta'(q_s, \epsilon) &= Q \\ \delta'(q_s, a) &= \emptyset & \forall a \in \Sigma \\ \delta'(q, b) &= \delta(q, b) & \forall q \in Q, b \in \Sigma \cup \{\epsilon\} \end{aligned}$$

### Hausaufgabe 4 (5 Punkte)

1. Seien  $A_1 = (Q_1, \Sigma, \delta_1, q_{01}, F_1)$  und  $A_2 = (Q_2, \Sigma, \delta_2, q_{02}, F_2)$  beliebige DFAs, und  $M = (Q_M, \Sigma, \delta_M, q_{0M}, F_M)$  ein minimaler DFA mit  $L(M) = L(A_1) \cup L(A_2)$ . Zeigen Sie:

$$|Q_M| \leq |Q_1| \cdot |Q_2|.$$

2. Berechnen Sie systematisch den Quotientenautomaten  $M/\equiv_M$  des DFA  $M = (\{A, B, C, D, E, F\}, \{0, 1\}, \delta, A, \{A, D\})$  mit folgender Übergangsfunktion:

$q$	$\delta(q, 0)$	$\delta(q, 1)$
$A$	$A$	$F$
$B$	$C$	$A$
$C$	$D$	$B$
$D$	$D$	$C$
$E$	$F$	$D$
$F$	$A$	$E$

Protokollieren Sie die Schritte des angewandten Verfahrens.

## Lösungsvorschlag

1. Der Produktautomat  $A_1 \times A_2$  von  $A_1$  und  $A_2$  mit  $|Q_1| \cdot |Q_2|$  Zuständen akzeptiert  $L(A_1) \cap L(A_2)$ .

Es gilt  $L(A_1) \cup L(A_2) = \overline{\overline{L(A_1)} \cap \overline{L(A_2)}}$ .

Da es zu jedem DFA  $A$  einen DFA  $A'$  mit gleichvielen Zuständen gibt, der  $\overline{L(A)}$  akzeptiert, gibt es einen DFA mit  $|Q_1| \cdot |Q_2|$  Zuständen, der  $L(M)$  akzeptiert.

2.

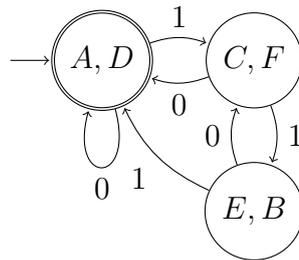
	A				
×	B				
×		C			
	×	×	D		
×			×	E	
×			×		F

(1) Nach der Initialisierung

	A				
×	B				
×	×	C			
	×	×	D		
×		×	×	E	
×	×		×	×	F

(2) Nach dem 2. Durchlauf

Ergebnis nach Kollabierung:



## Vorbereitung 1

Sei  $\rightarrow \subseteq A \times A$  eine beliebige binäre Relation. Dann sind die Relationen  $\rightarrow^n$ ,  $\rightarrow^*$  und  $\rightarrow^+$  wie in Def. 3.5 definiert.

1. Beweisen oder widerlegen Sie:

$$(a) \quad x \rightarrow y \iff x \rightarrow^1 y$$

$$(b) \quad x \rightarrow y \implies x \rightarrow^2 y$$

$$(c) \quad x \rightarrow y \implies x \rightarrow^* y$$

2. Zeigen Sie formal mit Induktion:  $x \rightarrow^n y \wedge y \rightarrow^m z \implies x \rightarrow^{n+m} z$

## Lösungsvorschlag

$$1. \quad (a) \quad x \rightarrow y \iff (\exists z. x = z \wedge z \rightarrow y) \iff (\exists z. x \rightarrow^0 z \wedge z \rightarrow y) \iff x \rightarrow^1 y$$

(b) Die Aussage ist im allgemeinen falsch: Als Gegenbeispiel wählen wir  $A = \mathbb{N}$  und  $\rightarrow = \{(n, m) \mid n < m\}$ . Dann gilt  $0 \rightarrow 1$ , aber nicht  $0 \rightarrow^2 1$ , denn es gibt kein  $z$  mit  $0 \rightarrow z \rightarrow 1$ .

$$(c) \quad x \rightarrow y \iff x \rightarrow^1 y \implies (\exists n. x \rightarrow^n y) \iff x \rightarrow^* y$$

2. Induktion über  $m$ : Wenn  $m = 0$ , dann ist  $x = y$  und es ist nichts zu zeigen. Im Induktionsschritt schließen wir wie folgt:

$$\begin{aligned} & x \rightarrow^n y \wedge y \rightarrow^{m+1} z \\ \implies & x \rightarrow^n y \wedge y \rightarrow^m w \wedge w \rightarrow z && \text{(nach Def.)} \\ \implies & x \rightarrow^{n+m} w \wedge w \rightarrow z && \text{(nach Induktionshypothese)} \\ \implies & x \rightarrow^{n+m+1} z && \text{(nach Def.)} \end{aligned}$$

## Vorbereitung 2

Gegeben sei die Grammatik  $G = (\{S, T\}, \{a, b\}, P, S)$  mit den Produktionen

$$S \rightarrow Tb, \quad T \rightarrow aS \mid Sb \mid \epsilon.$$

1. Beschreiben Sie  $L(G)$  mengentheoretisch möglichst einfach.

2. Beweisen oder widerlegen Sie: (i)  $a^2b^4 \in L(G)$ . (ii)  $a^4b^4 \in L(G)$ .

## Lösungsvorschlag

1.  $L(G) = \{a^m b^n \mid 0 \leq m < n \text{ und } m + n \text{ ungerade}\}$ .

Man beachte, dass nicht verlangt wurde, die Darstellung zu beweisen.

2. Es gilt  $a^2 b^4 \notin L(G)$  und  $a^4 b^4 \notin L(G)$ . Dies würde sofort aus der Teilaufgabe 1 folgen. Allerdings wurde die Mengendarstellung für  $L(G)$  dort nicht bewiesen.

Es genügt aber der Nachweis, dass alle Satzformen  $\alpha$  der Grammatik  $G$ , die eine Variable  $T$  enthalten, eine gerade Länge haben. Da jedes Wort  $w \in L(G)$  mit der Produktion  $T \rightarrow_G \epsilon$  aus einer Satzform  $\alpha$  abgeleitet wird, die genau ein  $T$  enthält, folgt dann, dass  $w$  ungerade Länge haben muss.

Wir betten unsere zu beweisende Aussage ein in die umfassendere Aussage: Jede aus  $S$  ableitbare Satzform enthält höchstens eine Variable und jede aus  $S$  ableitbare Satzform  $\alpha$ , die eine Variable  $T$  enthält, besitzt eine gerade Länge, Und jede aus  $S$  ableitbare Satzform  $\beta$ , die keine Variable  $T$  enthält, besitzt eine ungerade Länge.

Der Beweis wird induktiv über die Länge der aus  $S$  ableitbaren Satzformen geführt.

Für  $\alpha = S$  gilt die Aussage offensichtlich ( $S$  wird aus  $S$  in 0 Schritten abgeleitet).

Jede Satzform  $\beta$ , die genau ein  $S$  enthält wird durch Anwendung der Produktion  $S \rightarrow Tb$  in eine Satzform  $\alpha$  überführt, die genau ein  $T$  enthält und um 1 länger ist. Wenn die Aussage für  $\alpha$  gegolten hat, dann gilt sie auch für  $\beta$ .

Jede Satzform  $\alpha$ , die genau ein  $T$  enthält wird durch Anwendung einer der 3 möglichen Produktionen in eine Satzform  $\beta$  überführt, die genau ein  $S$  enthält und um 1 länger ist, oder die keine Variablen enthält und um 1 kürzer ist. Wenn die Aussage für  $\beta$  gegolten hat, dann gilt sie auch für  $\alpha$ .

## Tutoraufgabe 1

Sei  $G = (V, \Sigma, P, S)$  eine kontextfreie Grammatik.

Zeigen Sie, dass für alle  $\alpha, \beta, u, v, \alpha_1, \alpha_2, \beta_1, \beta_2 \in (\Sigma \cup V)^*$  gilt:

1.  $\alpha \rightarrow_G \beta \implies u\alpha v \rightarrow_G u\beta v$  (Abschluss von  $\rightarrow_G$  unter Kontext)

2.  $\alpha \rightarrow_G^n \beta \implies u\alpha v \rightarrow_G^n u\beta v$  (Abschluss von  $\rightarrow_G^n$  unter Kontext)

3.  $\alpha_1 \rightarrow_G^n \beta_1 \wedge \alpha_2 \rightarrow_G^m \beta_2 \implies \alpha_1 \alpha_2 \rightarrow_G^{n+m} \beta_1 \beta_2$   
(Dekompositionslemma (Lemma 3.9), 1. Teil:  $\Leftarrow$ )

## Lösungsvorschlag

1. Wenn  $\alpha \rightarrow_G \beta$ , dann gibt es nach Def. 3.4 eine Regel  $A \rightarrow \gamma \in P$  und  $\alpha_1, \alpha_2$ , so dass  $\alpha = \alpha_1 A \alpha_2$  und  $\beta = \alpha_1 \gamma \alpha_2$ . Nun setzen wir  $\alpha'_1 = u\alpha_1$  und  $\alpha'_2 = \alpha_2 v$ , womit wieder die Definition von  $\rightarrow_G$  erfüllt ist, denn  $u\alpha v = \alpha'_1 A \alpha'_2$  und  $u\beta v = \alpha'_1 \gamma \alpha'_2$ . Somit gilt  $u\alpha v \rightarrow_G u\beta v$ .

2. Die zuvor gezeigte Eigenschaft übertragen wir mit Induktion über  $n$  auf die Mehrschrittrelation  $\rightarrow_G^n$ :

$$\text{Fall } n = 0: \alpha \rightarrow_G^0 \beta \implies \alpha = \beta \implies u\alpha v = u\beta v \implies u\alpha v \rightarrow_G^0 u\beta v.$$

Fall  $n \rightarrow n + 1$ : Wenn  $\alpha \rightarrow_G^{n+1} \beta$ , dann

$$\alpha \rightarrow_G^n \gamma \rightarrow_G \beta \implies u\alpha v \rightarrow_G^n u\gamma v \rightarrow_G u\beta v \implies u\alpha v \rightarrow_G^{n+1} u\beta v$$

3. Wir setzen jetzt lediglich die zuvor gezeigten Eigenschaften zusammen. Mit dem Abschluss unter Kontext schließen wir:

$$\alpha_1 \rightarrow_G^n \beta_1 \implies \alpha_1 \alpha_2 \rightarrow_G^n \beta_1 \alpha_2$$

$$\alpha_2 \rightarrow_G^m \beta_2 \implies \beta_1 \alpha_2 \rightarrow_G^m \beta_1 \beta_2$$

Mit der Kompositionseigenschaft aus VA1.2 folgt nun  $\alpha_1 \alpha_2 \rightarrow_G^{n+m} \beta_1 \beta_2$

## Tutoraufgabe 2

Geben Sie eine kontextfreie Grammatik an, die die Sprache

$$L = \{w \in \{a, b\}^* \mid \#_a(w) = 2 \cdot \#_b(w)\}$$

erzeugt und zeigen Sie die Korrektheit Ihrer Konstruktion.

### Lösungsvorschlag

Es sei  $G = (\{S\}, \{a, b\}, P, S)$  mit den Produktionen:

$$S \rightarrow \epsilon \mid SaSaSbS \mid SaSbSaS \mid SbSaSaS.$$

An den Produktionen der Grammatik fällt auf, dass vor und nach jedem terminalen Zeichen die Variable  $S$  steht. Damit ist sofort klar, dass ohne die Anwendung von  $\epsilon$ -Produktionen sich diese Eigenschaft auf alle abgeleiteten Satzformen  $\alpha$  überträgt. Da die  $\epsilon$ -Produktionen mit allen anderen Produktionen vertauscht werden können gilt für jedes ableitbare Wort  $w \in L(G)$

$$S \rightarrow_G^* w_S \rightarrow_\epsilon^* w,$$

wobei  $w_S$  für  $w = x_1 x_2 \dots x_n$  mit  $x_i \in \Sigma$  definiert sei durch

$$w_S = Sx_1 Sx_2 S \dots Sx_n S,$$

und  $\rightarrow_\epsilon^*$  eine Ableitung mit ausschließlicher Anwendung von  $\epsilon$ -Produktionen bedeuten soll. Den einfachen, induktiv geführten Beweis dieser Eigenschaft verlegen wir in die Schlußbemerkung.

Für den Beweis von  $L(G) = L$  müssen wir die zwei Richtungen  $L(G) \subseteq L$  und  $L \subseteq L(G)$  beweisen.

$L(G) \subseteq L$ : Wir zeigen mit Induktion über die Erzeugung von  $w$ , dass für jedes aus  $S$  ableitbare Wort  $w$  gilt  $\#_a(w) = 2 \cdot \#_b(w)$ . Dabei erhalten wir je einen Fall für jede Produktion.

- $S \rightarrow \epsilon$ : Für  $w = \epsilon$  gilt  $\#_a(w) = \#_b(w) = 0$ , d. h.  $w \in L$ .

- $S \rightarrow SaSaSbS$ : Sei  $w \in L(S)aL(S)aL(S)bL(S)$ , d.h.  $w = w_1aw_2aw_3bw_4$  mit  $w_i \in L(S)$ . Nach Induktionshypothese gilt  $\#_a(w_i) = 2 \cdot \#_b(w_i)$  und somit folgt  $\#_a(w) = \#_a(w_1aw_2aw_3bw_4) = \#_a(w_1) + \#_a(w_2) + \#_a(w_3) + \#_a(w_4) + 2 = 2 \cdot (\#_b(w_1) + \#_b(w_2) + \#_b(w_3) + \#_b(w_4) + 1) = 2 \cdot \#_b(w)$ .

Es folgt  $w \in L$ .

- Analog für die übrigen Produktionen.

$L \subseteq L(G)$ : Wir zeigen, dass jedes Wort  $w$  mit  $\#_a(w) = 2\#_b(w)$  aus  $S$  ableitbar ist. Wir führen die Induktion über die Länge von  $w$ :

- $|w| = 0$ : Dann ist  $w = \epsilon$  und  $S \rightarrow \epsilon$ .
- Es gibt keine Wörter aus  $L$  der Länge 1 oder 2.
- Alle Wörter aus  $L$  der Länge 3 ( $aab, aba, baa$ ) sind direkt in 5 Schritten (davon 4 mal  $S \rightarrow \epsilon$ ) ableitbar.
- $|w| = n + 1 \geq 4$ : Nach Induktionsvoraussetzung sind alle kürzeren Wörter  $w' \in L$  ableitbar.

Wegen  $|w| \geq 4$  und  $w \in L$  gibt es eine Zerlegung  $w = uaabv$  oder eine Zerlegung  $w = ubaav$ . Wir betrachten nur den Fall  $w = uaabv$ .

Es gilt offenbar  $w' = uv \in L$  und nach Induktionsvoraussetzung ist  $w'$  herleitbar, d.h.  $w' \in L(G)$ . Es folgt

$$S \rightarrow_G^* (uv)_S \rightarrow_{\epsilon}^* uSv \rightarrow_G^* uaabv = w.$$

Man beachte das allgemeine Schema: Für die Richtung  $L(G) \subseteq L$  verwendet man Induktion über die Erzeugung eines Worts. Für die Gegenrichtung verwendet man Eigenschaften der Sprache  $L$  und (meist) ein anderes Induktionsprinzip, hier die Induktion über die Länge.

Schlussbemerkung: Wir zeigen noch durch Induktion für alle  $w \in L(G)$

$$S \rightarrow_G^* w_S \rightarrow_{\epsilon}^* w,$$

- $S \rightarrow \epsilon$ : Für  $w = \epsilon$  gilt  $S \rightarrow_{\epsilon}^* w$ .
- $S \rightarrow SaSaSbS$ : Sei  $w \in L(S)aL(S)aL(S)bL(S)$ , d.h.  $w = w_1aw_2aw_3bw_4$  mit  $w_i \in L(S)$ . Nach Induktionshypothese gilt  $S \rightarrow_G^* (w_i)_S \rightarrow_{\epsilon}^* (w_i)$ . Es folgt

$$S \rightarrow_G^* SaSaSbS \rightarrow_G^* (w_1)_S a (w_2)_S a (w_3)_S b (w_4)_S = w_S \rightarrow_{\epsilon}^* w.$$

- Analog für die übrigen Produktionen.