# 5. Reguläre Sprachen

Wir verschieben den Fokus von endl. Automaten auf die Klasse der von diesen erkannten Sprachen. Dabei spielen endl. Automaten weiterhin eine wichtige Rolle. Wegen Satz 4.11 beschränken wir uns auf det. endl. Autom.

Sei A DEA. Die Menge der zulässigen Eingaben $\Sigma^*$ ist unendlich groß, die Menge der Zustände Q ist aber endlich. Zwangsläufig wird A also das Einlesen verschiedener Eingaben im gleichen Zustand abschließen (und somit gleich behandeln). Dies führt zum Begriff der A-Äquivalenz.

## Definition 5.1 (Äquivalenzrelation)

Sei $A$ eine Menge. Eine Äquivalenzrelation auf $A$ ist eine Relation $\sim \subseteq A^2$, so dass die folgenden Eigenschaften erfüllt sind (wie bei Relationen üblich verwenden wir Infixnotation.)

(i) $a \sim a \quad \forall a \in A$ (Reflexivität)

(ii) $a \sim b \Rightarrow b \sim a \quad \forall a, b \in A$ (Symmetrie)

(iii) $a \sim b,\ b \sim c \Rightarrow a \sim c \quad \forall a, b, c \in A$ (Transitivität).

Die Äquivalenzklasse eines Elementes $a \in A$ bezüglich $\sim$ ist die Menge $[a]_\sim := \{a' \in A : a' \sim a\}$. Der Index von $\sim$ ist die Kardinalität der Menge $A/\sim := \{[a]_\sim : a \in A\}$ falls diese endl. ist und $\infty$ andernfalls.

## Definition 5.2 (A-Äquivalenz)

Sei $A = (Q, \Sigma, \Delta, s, F)$ ein DEA mit erw. Übergangsfkt $\delta^*: Q \times \Sigma \to Q$.

Die A-Äquivalenz ist die Relation $\sim_A$ auf $\Sigma^*$ mit

$$u \sim_A v \iff \delta^*(s,u) = \delta^*(s,v) \qquad \forall u,v \in \Sigma^*.$$

## Bemerkung 5.3

Sei $A = (Q, \Sigma, \Delta, s, F)$ ein DEA.

(i) Die A-Äquivalenz ist eine Äquivalenzrelation.

(ii) Der Index von $\sim_A$ ist höchstens $|Q|$.

(iii) Es gilt $L(A) = \bigcup_{w \in L(A)} [w]_{\sim_A}$.

## Definition 5.4 (Rechtskongruenz)

Sei $\Sigma$ ein Alphabet. Eine Rechtskongruenz auf $\Sigma^*$ ist eine Äquivalenzrelation $\sim \subseteq (\Sigma^*)^2$ mit $u \sim v \Rightarrow uw \sim vw \quad \forall u, v, w \in \Sigma^*$.

## Proposition 5.5

Sei $A = (Q, \Sigma, \Delta, s, F)$ ein DEA. Die $A$-Äquivalenz $\sim_A$ ist eine Rechtskongruenz auf $\Sigma^*$.

Beweis: Seien $u, v, w \in \Sigma^*$ mit $u \sim_A v$. Dann gilt

$$\delta^*_{det,A}(s, uw) = \delta^*_{det,A}(\delta^*_{det,A}(s,u), w) = \delta^*_{det,A}(\delta^*_{det,A}(s,v), w)$$

$$= \delta^*_{det,A}(s, vw).$$

(hier benutzen wir Bem 4.3 und 4.5)

Damit gilt $uw \sim_A vw$.

$\square$

Zu jedem DEA $A$ gibt es also eine dazugehörige Rechtskongruenz $\sim_A$ auf $\Sigma^*$ mit endl. Index, so dass $L(A)$ die Vereinigung von Äquivalenzklassen von $\sim_A$ ist. Tatsächlich gilt auch die Umkehrung: Ist $L$ die Vereinigung von Äquivalenzklassen einer Rechtskongr. $\sim$ mit endl. Index, so gibt es einen DEA $A$ mit $L(A) = L$.

Definition 5.6

Sei $\Sigma$ ein Alphabet und $L$ Vereinigung von Äquiv.-klassen einer Rechtskongr. $\sim$ auf $\Sigma^*$ mit endl. Index. Es bezeichne

$$A_{\sim, L} := \left( \Sigma^*/\sim, \Sigma, \Delta, [\lambda]_\sim, \{[w]_\sim : w \in L\} \right)$$

den DEA mit $\delta_{det, A_{\sim,L}}([w]_\sim, a) = [wa]_\sim \quad \forall w \in \Sigma^*$ und $a \in \Sigma$.

Die Wohldefiniertheit von $\delta_{det, A_{\sim,L}}$ ergibt sich daraus, dass $\sim$ eine Rechtskongr. ist.

Um uns davon zu überzeugen, dass $L(A_{\sim,L}) = L$ gilt betrachten wir zunächst die Arbeitsweise von $A_{\sim,L}$.

## Lemma 5.7

Sei $\Sigma$ ein Alphabet, $L$ Vereinigung von Äquiv.-klassen einer Rechtskongr. $\sim$ auf $\Sigma^*$ mit endl. Index und sei $\delta^* : \Sigma^*/_\sim \times \Sigma^* \to \Sigma^*/_\sim$ die erw. Übergangsfkt von $A_{\sim,L}$. Dann gilt $\delta^*([\lambda]_\sim, w) = [w]_\sim \; \forall w \in \Sigma^*$.

Beweis: Wir verwenden vollst. Ind. über $|w|$.

Es gilt $\delta^*([\lambda]_\sim, \lambda) = [\lambda]_\sim$.

Sei nun $w \in \Sigma^+$ mit $\delta^*([\lambda]_\sim, v) = [v]_\sim \; \forall v \in \Sigma^{\leq |w|-1}$. Es genügt zu zeigen, dass $\delta^*([\lambda]_\sim, w) = [w]_\sim$.

Sei $va = w$ mit $a \in \Sigma$.

Dann gilt $\delta^*([\lambda]_\sim, w) = \delta^*(\delta^*([\lambda]_\sim, v), a) = \delta^*([v]_\sim, a) = [va]_\sim = [w]_\sim$.

$\square$

## Satz 5.8

Sei $L$ die Vereinigung von Äquiv.-klassen einer Rechtskongr. $\sim$ mit endl. Index.

Es gilt $L(A_{\sim,L}) = L$.

<u>Beweis:</u> Sei $\Sigma$ das Alphabet, so dass $\sim$ eine Rechtskongr. auf $\Sigma^*$ ist.

Sei $\delta^*: \Sigma^*/_{\sim} \times \Sigma^* \longrightarrow \Sigma^*/_{\sim}$ die erw. Übergangsfkt

von $A_{\sim,L}$ und sei $w \in \Sigma^*$. Aus Lemma 5.7 folgt

$$w \in L(A_{\sim,L}) \iff \delta^*([\lambda]_\sim, w) \in \{[v]_\sim : v \in L\}$$

$$\iff [w]_\sim \in \{[v]_\sim : v \in L\}$$

$$\iff \exists v \in L: [w]_\sim = [v]_\sim$$

$$\iff \exists v \in L: w \sim v$$

$$\iff w \in L$$

$\square$

## Korollar 5.9

Eine Sprache L ist genau dann regulär, wenn sie die Vereinigung von Äquiv.-klassen einer Rechtskongr. mit endl. Index ist.

Beweis: Folgt aus Bem 5.3, Prop. 5.5 und Satz 5.8 □

Betrachtet man nur det. endl. Automaten ohne unerreichbare Zustände, so entsprechen diese bis auf Umbenennung von Zuständen sogar den Rechtskongr. mit endl. Index zusammen mit Vereinigungen von Äquivalenzklassen dieser.

## Definition 5.10 (erreichbar)

Sei $\Sigma$ ein Alphabet. Sei $A = (Q, \Sigma, \Delta, s, F)$ ein EA mit erw. Übergangsfkt $\delta^*$. Ein Zustand $q \in Q$ heißt erreichbar in A wenn es ein Wort $w \in \Sigma^*$ mit $q \in \delta^*(s, w)$ gibt.

## Definition 5.11 (isomorph)

Sei $A_i = (Q_i, \Sigma, \Delta_i, s_i, F_i)$ für $i \in \{1,2\}$ ein EA mit Übergangsfkt $\delta_i$.

Die endl. Automaten $A_1$ und $A_2$ sind isomorph, kurz $A_1 \cong A_2$, wenn es eine Bijektion $f: Q_1 \to Q_2$ gibt, so dass Folgendes gilt:

(i) $f(s_1) = s_2$

(ii) $\delta_2(f(q_1), a) = f(\delta_1(q_1, a)) \quad \forall q_1 \in Q_1, a \in \Sigma$

(iii) $f(F_1) = F_2$.

## Satz 5.12

(i) Ist $A$ ein DEA ohne unerreichbare Zustände, so gilt $A \cong A_{\sim_A, L(A)}$

(ii) Ist $L$ die Vereinigung von Äquivalenzklassen einer Rechtskongr. $\sim$ mit endl. Index, so gilt $(\sim, L) = (\sim_{A_{\sim,L}}, L(A_{\sim,L}))$.

<u>Beweis:</u> (i) Sei $A = (Q, \Sigma, \Delta, s, F)$ ein DEA mit erw. Übergangsfkt $\delta^*: Q \times \Sigma^* \to Q$ ohne unerreichbare Zustände, $\sim := \sim_A$, $A' := A_{\sim, L(A)}$ und sei $\delta': \Sigma^*/_\sim \times \Sigma^* \to \Sigma^*/_\sim$ die erw. Übergangsfkt von $A'$.

Sei $f: Q \to \Sigma^*/_\sim$ die Bij. mit $f(q) := \{w \in \Sigma^* : \delta^*(s,w) = q\}$,

Es gelten $f(s) = [\lambda]_\sim$ und $f(F) = \{[w]_\sim : w \in L(A)\}$.

Es genügt somit zu zeigen, dass $\delta'(f(q), a) = f(\delta(q, a)) \; \forall q \in Q, a \in \Sigma$.

Sei $q \in Q$, $a \in \Sigma$, $w \in \Sigma^*$.

Es genügt $w \in \delta'(f(q), a) \iff \delta^*(s, w) = \delta^*(q, a)$ zu zeigen.

Sei $v \in \Sigma^*$ mit $\delta^*(s, v) = q$. Nun gilt

$$w \in \delta'(f(q), a) \iff w \in \delta'([v]_\sim, a) \iff w \sim va$$

$$\iff \delta^*(s, w) = \delta^*(s, va) \iff \delta^*(s, w) = \delta^*(q, a)$$

Bem. 4.3 und 4.5

(ii) Sei $\Sigma$ ein Alphabet, $\sim$ eine Rechtskongr. auf $\Sigma^*$, $L$ Vereinigung von Äquiv.-klassen von $\sim$, $A' := A_{\sim, L} = (\Sigma^*/\sim, \Sigma, \Delta', [\lambda]_\sim, \eta)$, $\delta'^* : \Sigma^*/\sim \times \Sigma^* \to \Sigma^*/\sim$ die erw. Übergangsfkt von $A'$ und $\sim' := \sim_{A'}$.

$\{w \in \Sigma^* : w \in L\}$

Nach Satz 5.8 gilt $L = L(A')$, es genügt also $\sim = \sim'$ zu zeigen.

Seien $u, v \in \Sigma^*$. Aus Lemma 5.7 folgt

$$u \sim v \iff [u]_\sim = [v]_\sim \iff \delta'([\lambda]_\sim, u) = \delta'([\lambda]_\sim, v) \iff u \sim' v \qquad \square$$

Satz 5.12 bedeutet insbesondere folgendes:

Ist $A_i$, $i \in \{1,2\}$ ein DEA ohne unerreichbare Zustände, so gilt

$A_1 \cong A_2 \iff (\sim_{A_1}, L(A_1)) = (\sim_{A_2}, L(A_2))$ und ist $L_i$ für $i \in \{1,2\}$

Vereinigung von Äquivalenzklassen einer Rechtskongr. $\sim_i$ mit endl. Index,

so gilt $(\sim_1, L_1) = (\sim_2, L_2) \iff A_{\sim_1, L_1} \cong A_{\sim_2, L_2}$.

Ist $L$ eine reg. Sprache, so gibt es versch. endl. Automaten (ohne unerreichbare Zustände) mit $L(A) = L$. Äquivalent gilt:

$L$ ist die Vereinigung von Äquiv.-klassen verschiedener Rechtskongr. mit endl. Index. Für alle solche Rechtskongr. $\sim$ und $\forall u, v, w \in \Sigma^*$ mit

$u \sim v$ gilt aber

$$uw \in L \iff \delta^*_{det,A}(s,uw) \in F \iff \delta^*_{det,A}(s,vw) \in F \iff vw \in L$$

Dies führt zum Begriff der L-Äquivalenz und zeigt, dass die Partition in die Äquiv.-klassen von $\sim$ Verfeinerung der Part. in die Äquiv-klassen der L-Äquivalenz ist.

**Definition 5.13** (L-Äquivalenz)

Sei $L$ eine Sprache über einem Alphabet $\Sigma$. Die L-Äquivalenz von L als Sprache ist die Relation $\sim_L$ auf $\Sigma^*$ mit

$$u \sim_L v \iff \left( uw \in L \iff vw \in L \quad \forall w \in \Sigma^* \right).$$

## Bemerkung 5.14

Sei $L$ eine Sprache über $\Sigma$.

(i) Die $L$-Äquivalenz ist eine Rechtskongr.

(ii) Es gilt $L = \bigcup_{w \in L} [w]_{\sim_L}$.

## Definition 5.15 (Partition)

Sei $A$ eine Menge. Eine Partition von $A$ ist eine Menge $\mathcal{A} = \{A_1, \dots, A_n\}$ paarweise disj. nichtl. Teilmengen von $A$ mit $\bigcup_{i \in [n]} A_i = A$.

## Definition 5.16 (Verfeinerung)

Seien $\mathcal{A}_1$ und $\mathcal{A}_2$ Partitionen einer Menge $A$. Die Partition $\mathcal{A}_2$ verfeinert $\mathcal{A}_1$ (heißt Verfeinerung von $\mathcal{A}_1$), wenn es $\forall A_2 \in \mathcal{A}_2$ ein $A_1 \in \mathcal{A}_1$ mit $A_2 \subseteq A_1$ gibt.

## Bemerkung 5.17

Seien $\mathcal{A}_1$ und $\mathcal{A}_2$ Partitionen einer Menge $A$, so dass $\mathcal{A}_2$ die Part. $\mathcal{A}_1$ verfeinert.

(i) $\forall A' \in \mathcal{A}_1$, gibt es eine Teilmenge $\mathcal{A}_2' \subseteq \mathcal{A}_2$, die eine Partition von $A'$ ist.

(ii) Es gilt $|\mathcal{A}_1| \leq |\mathcal{A}_2|$

(iii) Gilt $|\mathcal{A}_1| = |\mathcal{A}_2|$ dann ist $\mathcal{A}_1 = \mathcal{A}_2$.

## Proposition 5.18

Sei $\Sigma$ ein Alphabet und $L$ eine Sprache über $\Sigma$ und $\sim$ eine Rechtskongr. auf $\Sigma^*$ mit $L = \bigcup_{w \in L} [w]_\sim$. Die Partition $\Sigma^*/_\sim$ ist eine Verfeinerung der Partition $\Sigma^*/_{\sim_L}$.

**Beweis:** Seien $u, v \in \Sigma^*$ mit $u \sim v$. Es genügt zu zeigen, dass $u \sim_L v$.

Sei $w \in \Sigma^*$. Es genügt $uw \in L \Leftrightarrow vw \in L$ zu zeigen.

Da $\sim$ eine Rechtskongr. ist folgt $uw \sim vw$. Ist $uw \in L$, so folgt aus $L = \bigcup_{w' \in L} [w']_\sim$ auch $vw \in L$. (analog folgt auch $vw \in L \Rightarrow uw \in L$).

$\Rightarrow u \sim_L v$. □

D.h. $\sim_L$ ist die gröbste Partition, die $L$ darstellen kann.

Definition 5.19 (Minimalautomat)

Sei $L$ eine reg. Sprache über $\Sigma$. Der Minimalautomat von $L$ als Sprache über $\Sigma$ ist der DEA $A_{\sim_L, L}$.

## Satz 5.20

Sei $L$ eine reg. Sprache über $\Sigma$ und sei $M = (Q, \Sigma, \delta, s, F)$ der Minimalautomat von $L$. Dann gilt:

(i) $L(M) = L$

(ii) Ist $A$ ein DEA mit Zustandsmenge $Q_A$ und $L(A) = L$, so gilt $|Q_A| \geq |Q|$.

(iii) Ist $A$ ein DEA mit $|Q|$ Zuständen und $L(A) = L$, so gilt $A \cong M$.

**Beweis:** (i) Folgt direkt aus Satz 5.8

(ii) Aus Bem 5.3 (ii) folgt $|Q_A| \geq |\Sigma^* /_{\sim_A}|$. Nach Prop. 5.18 ist $\Sigma^*/_{\sim_A}$ eine Verfeinerung von $\Sigma^*/_{\sim_L}$, nach Bem. 5.17 (ii) gilt also $|\Sigma^*/_{\sim_A}| \geq |\Sigma^*/_{\sim_L}|$.

Wegen $|\Sigma^*/_{\sim_L}| = |Q|$ folgt somit

$$|Q_A| \geq |\Sigma^*/_{\sim_A}| \geq |\Sigma^*/_{\sim_L}| = |Q|.$$

(iii) Sei $A$ ein DEA mit $|Q|$ Zuständen und $L(A) = L$.

Hätte $A$ unerreichbare Zustände, so folgt $|\Sigma^{*}/{\sim_A}| < |Q| = |\Sigma^{*}/{\sim_L}|$ im Widerspruch zu Bem 5.17 (ii) und Prop. 5.18.

Nach Satz 5.12 genügt es zu zeigen, dass $\sim_A = \sim_M$ zu zeigen.

Die Relationen $\sim_M$ und $\sim_A$ sind nach Bem 5.3 und Prop. 5.5 Rechtskongr. mit endlichen Index und $L$ ist Vereinigung von Äquivalenzklassen davon. Damit sind $\Sigma^{*}/{\sim_A}$ und $\Sigma^{*}/{\sim_M}$ nach Prop. 5.18 Verfeinerungen von $\Sigma^{*}/{\sim_L}$. Somit sind die Indices von $\sim_A$ und $\sim_M$ mind. so groß wie der Index von $\sim_L$. Weiter sind die Indices von $\sim_A$ und $\sim_M$ nach Bem. 5.3 (ii) aber auch höchstens so groß wie $|Q| = |\Sigma^{*}/{\sim_L}|$.

Die Indices von $\sim_A$, $\sim_M$ und $\sim_L$ sind alle gleich groß. Da $\Sigma^{*}/{\sim_A}$ und $\Sigma^{*}/{\sim_M}$ Verfeinerungen von $\Sigma^{*}/{\sim_L}$ sind, folgt mit Bem 5.17 (iii) somit

$$\Sigma^{*}/{\sim_A} = \Sigma^{*}/{\sim_M} = \Sigma^{*}/{\sim_L} \quad \text{und} \quad \sim_A = \sim_L = \sim_M.$$

$\square$

Unsere bisherigen Betrachtungen erlauben verschiedene äquivalenten Charakterisierungen der Klasse der reg. Sprachen.

Satz 5.21 (Satz von Myhill und Nerode)

Für eine Sprache $L$ über einem Alphabet $\Sigma$ sind die folgenden Aussagen äquivalent:

(i) $L$ ist regulär

(ii) Der Index von $\sim_L$ ist endlich.

(iii) $L$ ist die Vereinigung von Äqui-klassen einer Rechtskongr. mit endl. Index.

Beweis: (i) $\Leftrightarrow$ (ii) ist die Aussage von Kor. 5.9.

Die Relation $\sim_L$ ist nach Bem 5.14 eine Rechtskongruenz und es gilt $L = \bigcup_{w \in L} [w]_{\sim_L}$. Somit folgt (ii) $\Rightarrow$ (iii).

Die Implikation (iii) $\Rightarrow$ (ii) folgt aus Bem. 5.17 (ii) und Prop. 5.18

$\square$

Wir wollen nun ein Kriterium beschreiben das hilft nicht reg. Sprachen zu erkennen.

Satz 5.22 (Pumping-Lemma)

Sei $\Sigma$ ein Alphabet. Für jede reguläre Sprache $L \subseteq \Sigma^*$ gibt es eine Konstante $k \in \mathbb{N}$, so dass folgendes gilt:

Ist $z \in L$ mit $|z| \geq k$, so gibt es Wörter $u, v, w \in \Sigma^*$ mit $z = uvw$, so dass folgendes gilt:

(i) $v \neq \lambda$

(ii) $|uv| \leq k$

(iii) $uv^iw \in L \quad \forall i \in \mathbb{N}_0$.

Beweis: Sei $L \subseteq \Sigma^*$ eine reg. Sprache und $A = (Q, \Sigma, \delta, s, F)$ ein DFA mit $L(A) = L$ und erw. Übergangsfkt $\delta^* : Q \times \Sigma^* \to Q$.

Sei $k := |Q|$.

Sei $z \in L$ mit $|z| \geq k$. (Falls kein solches $z$ ex. ist nichts zu zeigen.)

Die Funktion $f: \{0, \ldots, k\} \to Q$, $i \mapsto \delta^*(s, z(1) \ldots z(i))$ ist keine Injektion, denn es gilt $|\{0, \ldots, k\}| = k+1 > |Q|$. Seien $j_1, j_2 \in \{0, \ldots, k\}$ mit $j_1 < j_2$ und $f(j_1) = f(j_2)$.

Sei $u := z(1) \ldots z(j_1)$, $v = z(j_1+1) \ldots z(j_2)$, $w = z(j_2+1) \ldots z(|z|)$.

Dann gilt $z = uvw$.

Aus $j_1 < j_2$ folgt $v \neq \lambda$.

Aus $j_2 \leq k$ folgt $|uv| \leq k$.

Es bleibt zu zeigen, dass $uv^i w \in L$ $\forall i \in \mathbb{N}_0$ gilt.

Dafür genügt es zu zeigen $\delta^*(s, uv^i) = \delta^*(s, u)$. $(*)$.

Denn dann gilt mit Bem 4.3 (iii) und 4.5

$$\delta^*(s, uv^i w) = \delta^*(\delta^*(s, uv^i), w) = \delta^*(\delta^*(s, u), w)$$

$$= \delta^*(\delta^*(s, uv), w) = \delta^*(s, uvw) \text{ und damit } uv^i w \in L.$$

Wir zeigen (*) mittels vollst. Ind. über $i$.

$i=0$ ✓

Gelte nun $\delta^*(s, uv^{i-1}) = \delta^*(s,u)$ für ein $i \in \mathbb{N}$.

Wieder folgt

$$\delta^*(s, uv^i) = \delta^*(\delta^*(s, uv^{i-1}), v) \overset{\text{I.H.}}{=} \delta^*(\delta^*(s,u), v)$$

$$= \delta^*(s, uv) = f(j_2) = f(j_1) = \delta^*(s,u).$$

□

Beispiel 5.23

Die Sprache $L = \{0^n 1^n : n \in \mathbb{N}_0\}$ ist nicht reg. Dies lässt sich mit dem Pumping-Lemma wie folgt zeigen.

Beweis: Ang. $L$ wäre reg.

$\overset{\text{PL}}{\Rightarrow}$ $\exists k \in \mathbb{N}_0 : \forall z \in L$ mit $|z| \geq k$ gilt, $\exists u, v, w \in \Sigma^*$ mit $z = uvw$ und

(i) $v \neq \lambda$ (ii) $|uv| \leq k$ (iii) $uv^iw \in L$ $\forall i \in \mathbb{N}_0$.

Sei $z := 0^k 1^k$.

Aus (i) und (ii) folgt, dass $v = 0^l$ für $l > 0$ und damit folgt

$uw = 0^{k-l} 1^k \in L$ nach PL. Dies ist ein Widerspruch, da

$L = \{ 0^n 1^n : n \in \mathbb{N}_0 \}$.

$\square$