Automaten und Formale Sprachen
4. Vorlesung

Prof. Dr. Dietrich Kuske

FG Automaten und Logik, TU Ilmenau

Wintersemester 2022/23

Proposition

Zu jeder rechtslinearen Grammatik $G = (V, \Sigma, P, S)$ gibt es einen NFA M mit $L(G) = L(M)$.

Beweis:
Wir definieren den NFA $M = (Z, \Sigma, S', \delta, E)$, wobei:

$$\begin{aligned} Z &= V \\ \delta(A, a) &= \{B \mid (A \to aB) \in P\} \\ S' &= \{S\} \\ E &= \{A \mid (A \to \varepsilon) \in P\} \end{aligned}$$

Behauptung: Für alle $w \in \Sigma^*$ und alle $A, B \in V$ gilt:

$$A \Rightarrow_G^* wB \iff B \in \widehat{\delta}\big(\{A\}, w\big)$$

Dies zeigt man per Induktion über $|w|$.

Sei nun $w \in \Sigma^*$.

Dann gilt:

$$\begin{aligned}
w \in L(G) &\iff \exists A \in V \colon S \Rightarrow_G^* wA \text{ und } (A \to \varepsilon) \in P \\
&\overset{\text{Beh.}}{\iff} \exists A \in V \colon A \in \widehat{\delta}\big(\{S\}, w\big) \text{ und } A \in E \\
&\iff \widehat{\delta}\big(\{S\}, w\big) \cap E \neq \emptyset \\
&\iff w \in L(M)
\end{aligned}$$

Es gilt also tatsächlich $L(G) = L(M)$. $\square$

NOTIZ (nach 4.3)

IA $w = \varepsilon$. Es gilt:

$$A \Rightarrow_G^* B \iff A = B \iff B \in \{A\} = \widehat{\delta}(\{A\}, \varepsilon)$$

IS Sei $w = av$ ($a \in \Sigma$, $v \in \Sigma^*$) und gelte die Behauptung für v.

$$\begin{aligned} A \Rightarrow_G^* avB \quad &\iff \quad \exists C \in V : (A \to aC) \in P \text{ und } C \Rightarrow_G^* vB \\ &\overset{\text{IV}}{\iff} \quad \exists C \in V : C \in \delta(A, a) \text{ und } B \in \widehat{\delta}(\{C\}, v) \\ &\iff \quad B \in \widehat{\delta}(\{A\}, av) \end{aligned}$$

NFAs, DFAs und rechtslineare Grammatiken

Satz

Seien Σ ein Alphabet und $L \subseteq \Sigma^*$ eine Sprache. Dann sind äquivalent:

1. L ist regulär (d.h. von einem DFA akzeptiert)
2. L wird von einem NFA akzeptiert
3. L ist rechtslinear (d.h. von einer Typ-3-Grammatik erzeugt)

verschiedene Modelle zur Beschreibung regulärer Sprachen:

- Rechtslineare Grammatiken: Verbindung zu Chomsky-Hierarchie, erzeugen Sprachen
  wenig geeignet für Entscheidung, ob Wort zu Sprache gehört
- NFAs: kompakte Darstellungen von Sprachen, intuitive graphische Darstellung
  wenig geeignet für Entscheidung, ob Wort zu Sprache gehört
- DFAs: u.U. exponentiell größer als NFA bzw. Grammatik
  gut geeignet für Entscheidung, ob Wort zu Sprache gehört

Abschlußeigenschaften

Definition

Gegeben sei eine Klasse K und ein n-stelliger Operator $\otimes\colon K^n \to K$.
Man sagt, eine Klasse $K' \subseteq K$ ist unter $\otimes$ abgeschlossen, wenn für beliebige Elemente $k_1, k_2, \ldots, k_n \in K'$ gilt: $\otimes(k_1, k_2, \ldots, k_n) \in K'$.

Wir betrachten hier Abschlußeigenschaften der Klasse aller regulären Sprachen (d.h. wir setzen $K =$ Klasse aller Sprachen und $K' =$ Klasse aller regulären Sprachen).

Uns interessiert dabei die folgende Frage:

> Falls L_1, L_2 regulär sind, sind dann auch $L_1 \cup L_2$, $L_1 \cap L_2$, $L_1 L_2$, $\overline{L_1} = \Sigma^* \backslash L_1$ (Komplement) und L_1^* regulär?

Kurze Antwort: Die regulären Sprachen sind unter all diesen Operationen abgeschlossen.

Satz

Wenn $L \subseteq \Sigma^*$ eine reguläre Sprache ist, dann ist auch $\Sigma^* \setminus L$ regulär.

Beweis: Da L regulär ist, gibt es einen DFA $M = (Z, \Sigma, z_0, \delta, E)$ mit $L(M) = L$. In diesem vertauschen wir die End- und Nicht-Endzustände, d.h. $M' = (Z, \Sigma, z_0, \delta, Z \setminus E)$.

Dann gilt für $w \in \Sigma^*$:

$$\begin{aligned} w \in \Sigma^* \setminus L &\iff w \notin L(M) \\ &\iff \widehat{\delta}(z_0, w) \notin E \\ &\iff \widehat{\delta}(z_0, w) \in Z \setminus E \\ &\iff w \in L(M') \,. \end{aligned}$$

Die Sprache $\Sigma^* \setminus L$ wird also von dem DFA M' akzeptiert und ist somit regulär. $\square$

Satz

Wenn L_1 und L_2 reguläre Sprachen sind, dann ist auch $L_1 \cup L_2$ regulär.

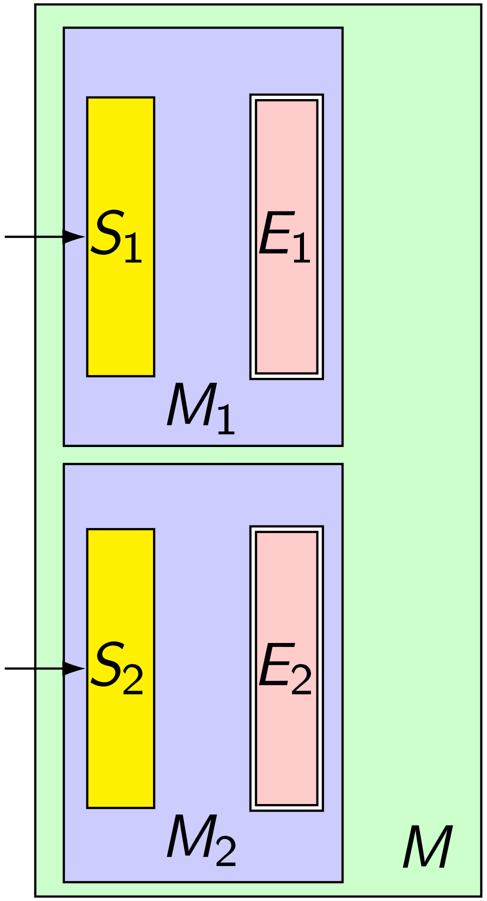
Beweis: Es gibt NFAs $M_i = (Z_i, \Sigma, S_i, \delta_i, E_i)$ für $i = 1, 2$ mit $L(M_i) = L_i$.
o.B.d.A. $Z_1 \cap Z_2 = \emptyset$.

Wir bauen nun einen NFA M als „disjunkte Vereinigung" von M_1 und M_2:

$$M = (Z_1 \cup Z_2, \Sigma, S_1 \cup S_2, \delta, E_1 \cup E_2)$$

wobei

$$\delta(z, a) = \begin{cases} \delta_1(z, a) & \text{für } z \in Z_1 \\ \delta_2(z, a) & \text{für } z \in Z_2 \end{cases}$$



Behauptung: Für alle $Y \subseteq Z_1 \cup Z_2$ und $w \in \Sigma^*$ gilt:

$$\widehat{\delta}(Y, w) = \widehat{\delta_1}(Y \cap Z_1, w) \cup \widehat{\delta_2}(Y \cap Z_2, w)$$

Wir zeigen diese Behauptung durch Induktion über $|w|$.

IA $w = \varepsilon$. Dann gilt

$$\widehat{\delta}(Y, \varepsilon) = Y = (Y \cap Z_1) \cup (Y \cap Z_2) = \widehat{\delta_1}(Y \cap Z_1, \varepsilon) \cup \widehat{\delta_2}(Y \cap Z_2, \varepsilon)$$

IS Sei $w = av$ $(a \in \Sigma, v \in \Sigma^*)$ und gelte die Behauptung für v.

$$\begin{aligned}
\widehat{\delta}(Y, av) &= \widehat{\delta}\big(\delta(Y, a), v\big) \\
&= \widehat{\delta}\Big(\underbrace{\delta_1(Y \cap Z_1, a)}_{\subseteq Z_1} \cup \underbrace{\delta_2(Y \cap Z_2, a)}_{\subseteq Z_2}, v\Big) \\
&\overset{IV}{=} \widehat{\delta_1}\big(\delta_1(Y \cap Z_1, a), v\big) \cup \widehat{\delta_2}\big(\delta_2(Y \cap Z_2, a), v\big) \\
&= \widehat{\delta_1}(Y \cap Z_1, av) \cup \widehat{\delta_2}(Y \cap Z_2, av)
\end{aligned}$$

Damit erhalten wir für $w \in \Sigma^*$:

$$
\begin{aligned}
\widehat{\delta}(S_1 \cup S_2, w) \cap (E_1 \cup E_2) &= \big(\underbrace{\widehat{\delta_1}(S_1, w)}_{\subseteq Z_1} \cup \underbrace{\widehat{\delta_2}(S_2, w)}_{\subseteq Z_2}\big) \cap (E_1 \cup E_2) \\
&= \big(\widehat{\delta_1}(S_1, w) \cap E_1\big) \cup \big(\widehat{\delta_2}(S_2, w) \cap E_2\big)
\end{aligned}
$$

und also

$$
\begin{aligned}
w \in L(M) &\iff \widehat{\delta}(S_1 \cup S_2, w) \cap (E_1 \cup E_2) \neq \emptyset \\
&\iff \widehat{\delta_1}(S_1, w) \cap E_1 \neq \emptyset \text{ oder } \widehat{\delta_2}(S_2, w) \cap E_2 \neq \emptyset \\
&\iff w \in L(M_1) \text{ oder } w \in L(M_2) \\
&\iff w \in L_1 \cup L_2
\end{aligned}
$$

Also wird die Sprache $L_1 \cup L_2$ von einem NFA akzeptiert und ist nach der Proposition auf Folie 3.22 damit regulär. $\square$

Satz

Wenn L_1 und L_2 reguläre Sprachen sind, dann ist auch $L_1 \cap L_2$ regulär.

Beweis:

Es gilt $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$ und wir wissen bereits, daß die Klasse der regulären Sprachen unter Komplement und Vereinigung abgeschlossen ist. $\square$

Bemerkung: Wird L_i von dem NFA $M_i = (Z_i, \dots)$ akzeptiert, so existiert also ein DFA (und damit ein NFA) mit $2^{2^{|Z_1|}+2^{|Z_2|}}$ Zuständen, der $L(M_1) \cap L(M_2)$ akzeptiert. Wir geben jetzt eine Konstruktion an, die mit $|Z_1| \cdot |Z_2|$ Zuständen auskommt (ohne Beweis).

Konstruktion:

In dieser Konstruktion werden die zwei Automaten für L_1 und L_2 miteinander synchronisiert und quasi „parallelgeschaltet“. Dies erfolgt durch das Bilden des Kreuzprodukts.

Seien $M_i = (Z_i, \Sigma, S_i, \delta_i, E_i)$ NFAs mit $L(M_i) = L_i$. Wir betrachten den NFA

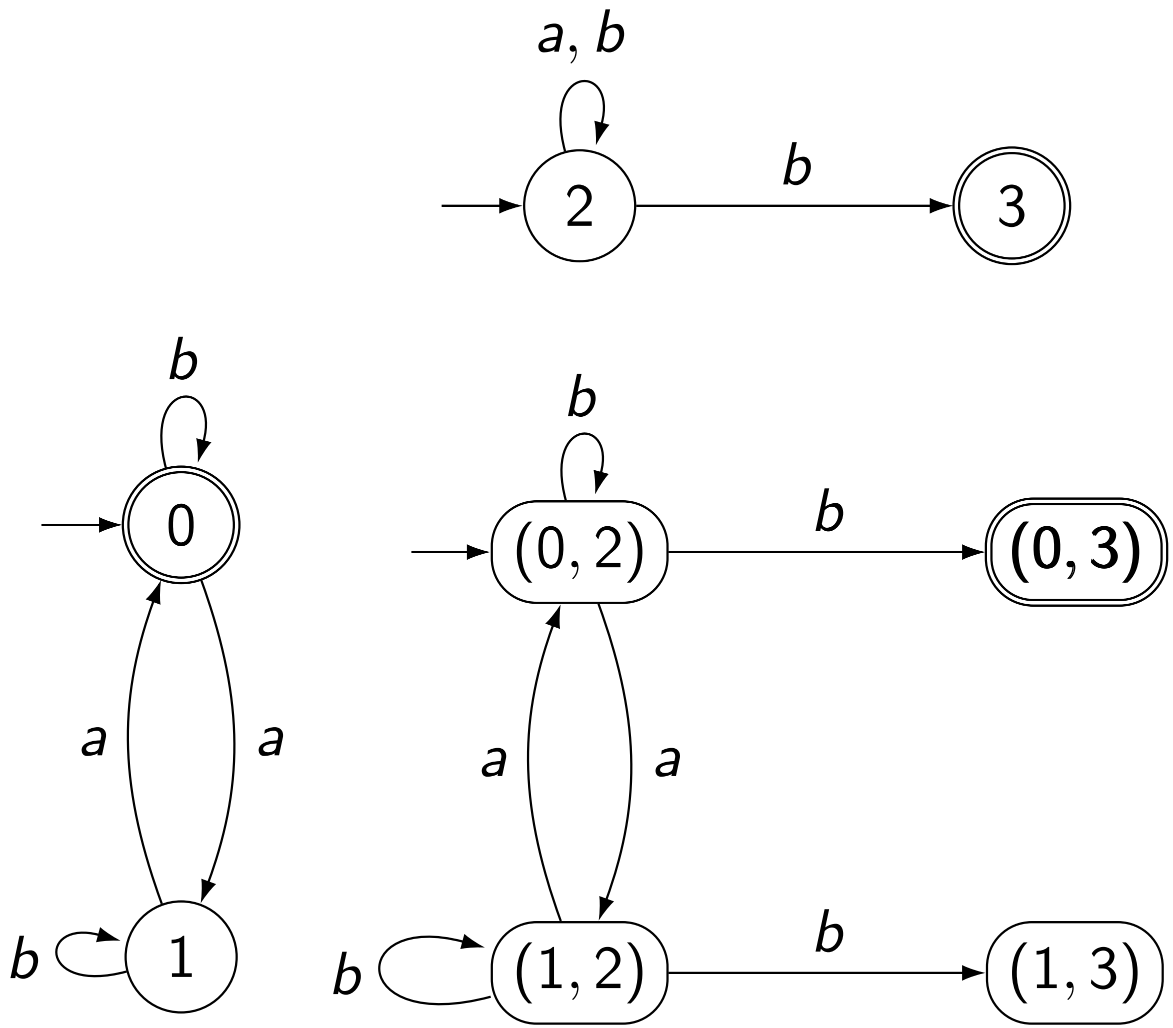
$$M = (\underbrace{Z_1 \times Z_2}_{Z}, \Sigma, \underbrace{S_1 \times S_2}_{S}, \delta, \underbrace{E_1 \times E_2}_{E}),$$

mit

$$\begin{aligned} \delta\big((z_1, z_2), a\big) &= \big\{(z_1', z_2') \mid z_1' \in \delta_1(z_1, a), z_2' \in \delta_2(z_2, a)\big\} \\ &= \delta_1(z_1, a) \times \delta_2(z_2, a)\,. \end{aligned}$$

Man nennt M auch das „Kreuzprodukt“ von M_1 und M_2.

Bildung des Kreuzprodukts zweier NFAs:



Behauptung: Für alle $Y \subseteq Z_1 \times Z_2$ und alle $w \in \Sigma^*$ gilt:

$$\begin{aligned}
\widehat{\delta}(Y, w) &= \bigcup_{(z_1,z_2)\in Y} \left(\widehat{\delta_1}\big(\{z_1\}, w\big) \times \widehat{\delta_2}\big(\{z_2\}, w\big)\right) \\
&= \left\{(z_1', z_2') \mid \exists (z_1, z_2) \in Y \colon z_1' \in \widehat{\delta_1}\big(\{z_1\}, w\big), z_2' \in \widehat{\delta_2}\big(\{z_2\}, w\big)\right\}
\end{aligned}$$

(der Beweis erfolgt per Induktion über $|w|$).

Konsequenz: Das „Kreuzprodukt“ von M_1 und M_2 akzeptiert die Sprache $L(M_1) \cap L(M_2)$.

Satz

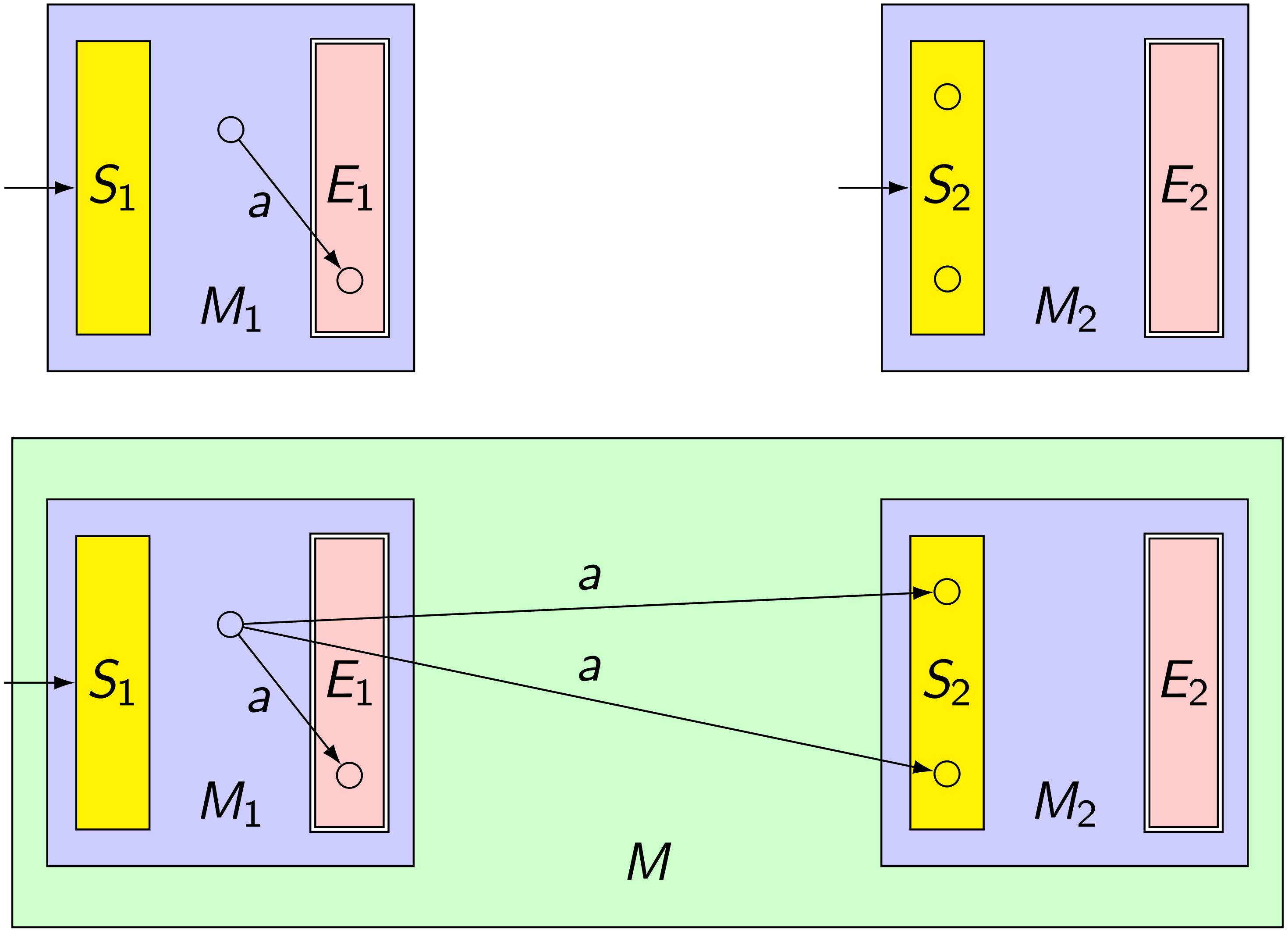
Wenn L_1 und L_2 reguläre Sprachen sind, dann ist auch $L_1 L_2$ regulär.

Beweis: Es gibt NFAs $M_i = (Z_i, \Sigma, S_i, \delta_i, E_i)$ mit $L(M_i) = L_i$ für $i = 1, 2$. o.B.d.A. $Z_1 \cap Z_2 = \emptyset$.

Wir verknüpfen nun M_1 und M_2 sequentiell zu einem NFA $M = (Z_1 \cup Z_2, \Sigma, S, \delta, E_2)$, wobei

$$S = \begin{cases} S_1 & \text{falls } \varepsilon \notin L_1 \\ S_1 \cup S_2 & \text{falls } \varepsilon \in L_1 \end{cases}$$

$$\delta(z, a) = \begin{cases} \delta_2(z, a) & \text{für } z \in Z_2 \\ \delta_1(z, a) & \text{für } z \in Z_1 \text{ mit } \delta_1(z, a) \cap E_1 = \emptyset \\ \delta_1(z, a) \cup S_2 & \text{für } z \in Z_1 \text{ mit } \delta_1(z, a) \cap E_1 \neq \emptyset \end{cases}$$



S_1
a
E_1
M_1
S_2
E_2
M_2
S_1
a
E_1
M_1
a
a
S_2
E_2
M_2
M

Behauptung: Für alle $Y \subseteq Z_1 \cup Z_2$ und $w \in \Sigma^+$ gilt:

$$\widehat{\delta}(Y, w) = \widehat{\delta_1}(Y \cap Z_1, w) \cup \widehat{\delta_2}(Y \cap Z_2, w) \cup \bigcup_{\substack{w=uv, u \neq \varepsilon \\ \widehat{\delta_1}(Y \cap Z_1, u) \cap E_1 \neq \emptyset}} \widehat{\delta_2}(S_2, v)$$

Wir zeigen diese Behauptung durch Induktion über $|w|$:

IA Sei $a \in \Sigma$. Dann gilt

$$\begin{aligned}
\widehat{\delta}(Y, a) &= \bigcup_{z \in Y} \delta(z, a) \\
&= \bigcup_{z \in Y \cap Z_1} \delta_1(z, a) \cup \bigcup_{z \in Y \cap Z_2} \delta_2(z, a) \\
&\qquad \cup \begin{cases} S_2 & \text{falls } \exists z \in Y \cap Z_1 : \delta_1(z, a) \cap E_1 \neq \emptyset \\ \emptyset & \text{sonst} \end{cases} \\
&= \widehat{\delta_1}(Y \cap Z_1, a) \cup \widehat{\delta_2}(Y \cap Z_2, a) \cup \bigcup_{\widehat{\delta_1}(Y \cap Z_1, a) \cap E_1 \neq \emptyset} \widehat{\delta_2}(S_2, \varepsilon)
\end{aligned}$$

IS Es ist $\widehat{\delta}(Y, aw)$ gleich

$$
\begin{aligned}
&= \widehat{\delta}\Big(\overbrace{\widehat{\delta_1}(Y \cap Z_1, a) \cup \widehat{\delta_2}(Y \cap Z_2, a) \cup \bigcup_{\widehat{\delta_1}(Y \cap Z_1, a) \cap E_1 \neq \emptyset} \widehat{\delta_2}(S_2, \varepsilon)}^{=:H}, w\Big) \\
&\overset{\text{IV}}{=} \widehat{\delta_1}(H \cap Z_1, w) \cup \widehat{\delta_2}(H \cap Z_2, w) \cup \bigcup_{\substack{w = uv, u \neq \varepsilon \\ \widehat{\delta_1}(H \cap Z_1, u) \cap E_1 \neq \emptyset}} \widehat{\delta_2}(S_2, v) \\
&= \widehat{\delta_1}(Y \cap Z_1, aw) \cup \widehat{\delta_2}(Y \cap Z_2, aw) \\
&\qquad \cup \widehat{\delta_2}\Big(\bigcup_{\widehat{\delta_1}(Y \cap Z_1, a) \cap E_1 \neq \emptyset} S_2, w\Big) \cup \bigcup_{\substack{w = uv, u \neq \varepsilon \\ \widehat{\delta_1}(Y \cap Z_1, au) \cap E_1 \neq \emptyset}} \widehat{\delta_2}(S_2, v)
\end{aligned}
$$

Damit ist die Behauptung bewiesen.

Für $w \in \Sigma^+$ erhalten wir also

$$
\begin{aligned}
w \in L(M) \quad &\iff \quad \widehat{\delta}(S, w) \cap E \neq \emptyset \\
&\iff \quad \widehat{\delta_2}(S \cap S_2, w) \cap E_2 \neq \emptyset \\
&\qquad\quad \text{oder } w = uv, u \neq \varepsilon, \widehat{\delta_1}(S \cap S_1, u) \cap E_1 \neq \emptyset, \\
&\qquad\qquad\quad \widehat{\delta_2}(S_2, v) \cap E_2 \neq \emptyset \\
&\iff \quad w = uv, u \in L(M_1), v \in L(M_2) \\
&\iff \quad w \in L_1 \cdot L_2 \,.
\end{aligned}
$$

Für $w = \varepsilon$ erhalten wir

$$\begin{aligned}
\varepsilon \in L(M) \quad &\iff \quad S \cap E_2 \neq \emptyset \\
&\iff \quad S_1 \cap E_2 \neq \emptyset \text{ oder } \big(\varepsilon \in L(M_1) \text{ und } S_2 \cap E_2 \neq \emptyset\big) \\
&\iff \quad \varepsilon \in L(M_1) \text{ und } \varepsilon \in L(M_2) \\
&\iff \quad \varepsilon \in L(M_1) \cdot L(M_2) = L_1 \cdot L_2 \,.
\end{aligned}$$

Also ist $L_1 \cdot L_2 = L(M)$ und damit nach der Proposition auf Folie 3.22 tatsächlich regulär. $\square$

Zusammenfassung 4. Vorlesung

in dieser Vorlesung neu

- NFAs, DFAs und rechtslineare Grammatiken sind tatsächlich gleich ausdrucksstark
- Abschluß der Klasse der rechtslinearen Sprachen unter Komplement, Vereinigung, Schnitt und Konkatenation

kommende Vorlesung

- Abschluß der Klasse der rechtslinearen Sprachen unter positiver Iteration und unter Kleene-Iteration
- alternative Beschreibung der rechtslinearen Sprachen mittels regulärer Ausdrücke