

FORMALE SYSTEME

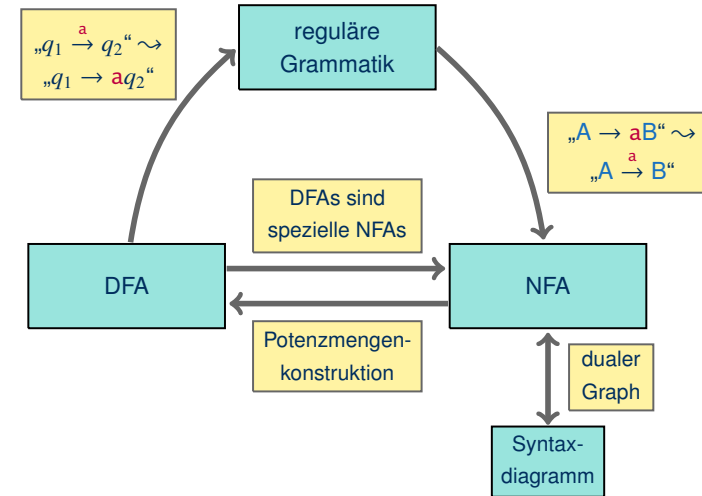
5. Vorlesung: Abschlusseigenschaften

Hannes Straß

Folien: © Markus Krötzsch, <https://iccl.inf.tu-dresden.de/web/FS2020>, CC BY 3.0 DE

TU Dresden, 25. Oktober 2021

Darstellungen von Typ-3-Sprachen



Hannes Straß, TU Dresden

Formale Systeme, VL 5

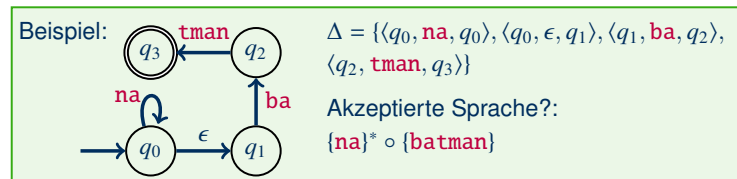
Folie 3 von 33

Automaten mit Wortübergängen

Ein nichtdeterministischer endlicher Automat mit Wortübergängen \mathcal{M} ist ein Tupel $\mathcal{M} = \langle Q, \Sigma, \Delta, Q_0, F \rangle$ mit den folgenden Teilen:

- Q : endliche Menge von Zuständen
- Σ : Alphabet
- Δ : Übergangsrelation, eine endliche Relation $\Delta \subseteq Q \times \Sigma^* \times Q$
- Q_0 : Menge möglicher Startzustände $Q_0 \subseteq Q$
- F : Menge von Endzuständen $F \subseteq Q$

Die Sprache solcher Automaten wird wie bei NFAs definiert, nur dass in einem Schritt beliebige Wörter gelesen werden können.



Hannes Straß, TU Dresden

Formale Systeme, VL 5

Folie 5 von 33

Wortübergänge machen NFAs nicht stärker

Satz: Jede von einem NFA mit Wortübergängen akzeptierte Sprache wird auch von einem „normalen“ NFA akzeptiert.

Beweis: Wir eliminieren zunächst Wortübergänge für Wörter $w \neq \epsilon$.

Für jeden Übergang $q \xrightarrow{w} q'$ mit $w = a_1 \dots a_n$ und $n \geq 2$:

- Führe $n - 1$ neue Zustände p_1, \dots, p_{n-1} ein.
- Ersetze $q \xrightarrow{w} q'$ durch Übergänge $q \xrightarrow{a_1} p_1, p_1 \xrightarrow{a_2} p_2, \dots, p_{n-1} \xrightarrow{a_n} q'$.

Es ist leicht zu sehen, dass diese Umformung korrekt ist.

Dadurch erhalten wir einen NFA, in dem nur noch zwei Formen von Übergängen vorkommen: $q_1 \xrightarrow{a} q_2$ und $q_1 \xrightarrow{\epsilon} q_2$.

Hannes Straß, TU Dresden

Formale Systeme, VL 5

Folie 6 von 33

ε-Übergänge machen NFA nicht stärker

Ein NFA mit Wortübergängen, bei dem alle Übergänge die Form $q_1 \xrightarrow{a} q_2$ oder $q_1 \xrightarrow{\epsilon} q_2$ haben, wird **ε-NFA** genannt.

Um den vorigen Beweis abzuschließen, müssen wir noch zeigen:

Satz: Jede von einem ε-NFA akzeptierte Sprache wird auch von einem „normalen“ NFA akzeptiert.

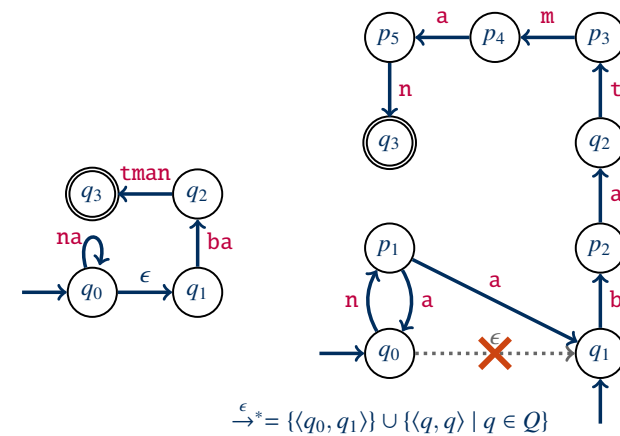
Beweis: Sei $\xrightarrow{\epsilon^*}$ der reflexive, transitive Abschluss von $\xrightarrow{\epsilon}$, d.h. die Menge aller Zustandspaare $\langle q, q' \rangle \in Q \times Q$, für die es Übergänge $q = p_0 \xrightarrow{\epsilon} p_1 \xrightarrow{\epsilon} \dots \xrightarrow{\epsilon} p_n = q'$ gibt, wobei $n \geq 0$.

Für einen ε-NFA $\mathcal{M} = \langle Q, \Sigma, \Delta, Q_0, F \rangle$ definieren wir einen NFA $\mathcal{M}' = \langle Q, \Sigma, \delta, Q'_0, F \rangle$ mit:

- $\delta(q, a) = \{q' \in Q \mid q \xrightarrow{a} r \xrightarrow{\epsilon^*} q' \text{ für ein } r \in Q\}$
- $Q'_0 = \{q \in Q \mid q_0 \xrightarrow{\epsilon^*} q \text{ für ein } q_0 \in Q_0\}$

Wir behaupten $\mathbf{L}(\mathcal{M}) = \mathbf{L}(\mathcal{M}')$.

Beispiel: Eliminierung von Wortübergängen



Korrektheit ε-NFA \rightsquigarrow NFA (1)

Wir behaupten $\mathbf{L}(\mathcal{M}) = \mathbf{L}(\mathcal{M}')$.

Beweis: Richtung $\mathbf{L}(\mathcal{M}) \supseteq \mathbf{L}(\mathcal{M}')$.

- Sei $w = a_1 \dots a_n \in \mathbf{L}(\mathcal{M}')$ und $p_0 p_1 \dots p_n$ ein akzeptierender Lauf für w in \mathcal{M}' .
- Dann gibt es in \mathcal{M}' Übergänge:

$$p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} p_n$$

- Dann gilt in \mathcal{M} :

$$q_0 \xrightarrow{\epsilon^*} p_0 \xrightarrow{a_1} q_1 \xrightarrow{\epsilon^*} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} q_n \xrightarrow{\epsilon^*} p_n$$

- Es ist $q_0 \in Q_0$ (da $p_0 \in Q'_0$) und $p_n \in F$, also hat \mathcal{M} einen akzeptierenden Lauf der Form $q_0 \dots p_0 q_1 \dots p_1 \dots q_n \dots p_n$.

Also ist $w \in \mathbf{L}(\mathcal{M})$.

Korrektheit ε-NFA \rightsquigarrow NFA (2)

Wir behaupten $\mathbf{L}(\mathcal{M}) = \mathbf{L}(\mathcal{M}')$.

Beweis: Richtung $\mathbf{L}(\mathcal{M}) \subseteq \mathbf{L}(\mathcal{M}')$.

- Sei $w = a_1 \dots a_n \in \mathbf{L}(\mathcal{M})$.
- Dann gibt es einen akzeptierenden Lauf mit Übergängen der Form

$$q_0 \xrightarrow{\epsilon^*} p_0 \xrightarrow{a_1} q_1 \xrightarrow{\epsilon^*} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} q_n \xrightarrow{\epsilon^*} p_n$$

- Dann gibt es in \mathcal{M}' Übergänge:

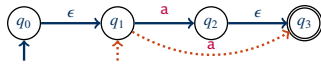
$$p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} p_n$$

- Dies führt zu einem akzeptierenden Lauf in \mathcal{M}' .

Also ist $w \in \mathbf{L}(\mathcal{M}')$. □

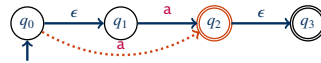
ϵ -NFA: Variationen

Die Konstruktion im Beweis „verlängert“ normale Übergänge „nach rechts“ durch Anhängen von ϵ -Transitionen.

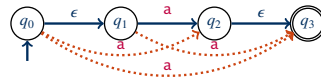


Man kann alternative Konstruktionen für den Beweis angeben:

- „Verlängerung nach links“: ϵ -Transitionen vor normalen Übergängen; Anfangszustände werden beibehalten; dafür werden Endzustände mit ϵ -Transitionen erweitert.



- „Verlängerung in beide Richtungen“: ϵ -Transitionen vor und nach normalen Übergängen; Anfangs- und Endzustände können beibehalten werden.*



*) Ausnahme: Anfangszustände mit ϵ -Pfad zu einem Endzustand werden Endzustand.

Quiz: NFA mit Wortübergängen

Quiz: Wir betrachten den NFA mit Wortübergängen $\mathcal{M} = \langle Q, \Sigma, \Delta, Q_0, F \rangle$ für $\Sigma = \{a, b\}$:

...

Abschlusseigenschaften

Abschlusseigenschaften regulärer Sprachen

Reguläre Sprachen sind unter zahlreichen Operationen abgeschlossen:

Satz: Wenn L , L_1 und L_2 reguläre Sprachen sind, dann beschreiben auch die folgenden Ausdrücke reguläre Sprachen:

- (1) $L_1 \cup L_2$ (Abschluss unter Vereinigung)
- (2) $L_1 \cap L_2$ (Abschluss unter Schnitt)
- (3) \bar{L} (Abschluss unter Komplement)
- (4) $L_1 \circ L_2$ (Abschluss unter Konkatenation)
- (5) L^* (Abschluss unter Kleene-Stern)

Beweisidee: Für jede Operation auf Sprachen entwickeln wir eine entsprechende *Operation auf Automaten*. Dadurch konstruieren wir Automaten, welche die gesuchte Sprache erkennen (daher muss die Sprache regulär sein).

Vereinigung von NFAs

Ein NFA zur Vereinigung von zwei NFAs ist leicht zu finden:

Gegeben seien zwei NFAs $\mathcal{M}_1 = \langle Q_1, \Sigma, \delta_1, Q_{0,1}, F_1 \rangle$ und $\mathcal{M}_2 = \langle Q_2, \Sigma, \delta_2, Q_{0,2}, F_2 \rangle$ mit $Q_1 \cap Q_2 = \emptyset$ (o.B.d.A.).

Der Vereinigungsautomat $\mathcal{M}_1 \oplus \mathcal{M}_2$ ist gegeben durch $\langle Q_1 \cup Q_2, \Sigma, \delta_{12}, Q_{0,1} \cup Q_{0,2}, F_1 \cup F_2 \rangle$, wobei

$$\delta_{12}(q, a) = \begin{cases} \delta_1(q, a) & \text{falls } q \in Q_1, \\ \delta_2(q, a) & \text{falls } q \in Q_2. \end{cases}$$

Das folgende Ergebnis ist leicht zu sehen:

Satz: $L(\mathcal{M}_1 \oplus \mathcal{M}_2) = L(\mathcal{M}_1) \cup L(\mathcal{M}_2)$.

Also sind reguläre Sprachen unter Vereinigung abgeschlossen.

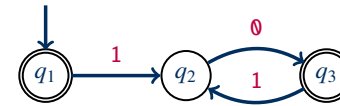
Abschluss unter Schnitt

Satz: Wenn L , L_1 und L_2 reguläre Sprachen sind, dann beschreiben auch die folgenden Ausdrücke reguläre Sprachen:

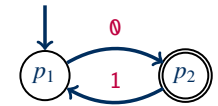
- (1) $L_1 \cup L_2$ (Abschluss unter Vereinigung)
- (2) $L_1 \cap L_2$ **Abschluss unter Schnitt**
- (3) \bar{L} (Abschluss unter Komplement)
- (4) $L_1 \circ L_2$ (Abschluss unter Konkatenation)
- (5) L^* (Abschluss unter Kleene-Stern)

Beispiel

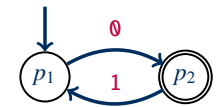
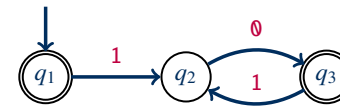
Die Vereinigung der NFAs



und



ergibt den NFA ...



Akzeptierte Sprache: $\{10\}^* \cup (\{01\}^* \circ \{0\}) = \{\epsilon, 0\} \circ \{10\}^*$

Schnitte mit Automaten

Gegeben: Automaten \mathcal{M}_1 und \mathcal{M}_2

Gesucht: Automat \mathcal{M} mit $L(\mathcal{M}) = L(\mathcal{M}_1) \cap L(\mathcal{M}_2)$

Idee:

- Verarbeite Eingabe gleichzeitig synchron auf beiden Automaten
- Akzeptiere genau dann, wenn beide Automaten akzeptieren

→ Produktautomat

Der Produktautomat

Erinnerung: Für gegebene Mengen A und B bezeichnet $A \times B$ die Menge aller möglichen Paare von Elementen aus A und B , d.h. $A \times B = \{\langle a, b \rangle \mid a \in A, b \in B\}$.

Gegeben seien zwei NFAs $\mathcal{M}_1 = \langle Q_1, \Sigma, \delta_1, Q_{0,1}, F_1 \rangle$ und $\mathcal{M}_2 = \langle Q_2, \Sigma, \delta_2, Q_{0,2}, F_2 \rangle$.

Der **Produktautomat** $\mathcal{M}_1 \otimes \mathcal{M}_2$ ist gegeben durch $\langle Q_1 \times Q_2, \Sigma, \delta, Q_{0,1} \times Q_{0,2}, F_1 \times F_2 \rangle$ wobei

$$\delta(\langle q_1, q_2 \rangle, a) = \delta_1(q_1, a) \times \delta_2(q_2, a)$$

Wir werden zeigen, dass dies der gesuchte Automat ist:

Satz: $L(\mathcal{M}_1 \otimes \mathcal{M}_2) = L(\mathcal{M}_1) \cap L(\mathcal{M}_2)$.

Also sind reguläre Sprachen unter Schnitt abgeschlossen.

Beobachtung: Wenn $|A| = |B| = 1$ gilt, dann ist $|A \times B| = 1$. Also ist der Produktautomat von DFAs ebenfalls ein DFA.

Beweis der Korrektheit des Produktautomaten

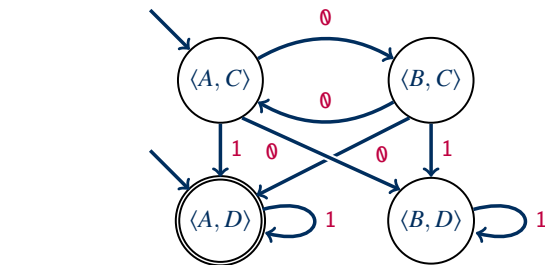
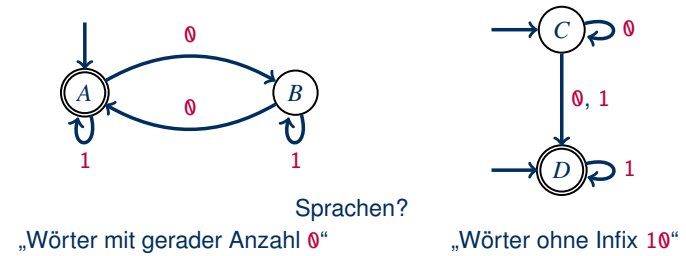
„ $L(\mathcal{M}_1 \otimes \mathcal{M}_2) \subseteq L(\mathcal{M}_1) \cap L(\mathcal{M}_2)$ “: Sei $w = a_1 \dots a_n \in L(\mathcal{M}_1 \otimes \mathcal{M}_2)$.

- Also gibt es einen akzeptierenden Lauf $\langle q_1^0, q_2^0 \rangle \langle q_1^1, q_2^1 \rangle \dots \langle q_1^n, q_2^n \rangle$
- Dann gilt:
 - $\langle q_1^0, q_2^0 \rangle \in Q_{0,1} \times Q_{0,2}$, also $q_1^0 \in Q_{0,1}$ und $q_2^0 \in Q_{0,2}$
 - $\langle q_1^n, q_2^n \rangle \in F_1 \times F_2$, also $q_1^n \in F_1$ und $q_2^n \in F_2$
 - $\langle q_1^i, q_2^i \rangle \in \delta(\langle q_1^{i-1}, q_2^{i-1} \rangle, a_i)$,
also $q_1^i \in \delta_1(q_1^{i-1}, a_i)$ und $q_2^i \in \delta_2(q_2^{i-1}, a_i)$
- Daher sind $q_1^0 q_1^1 \dots q_1^n$ und $q_2^0 q_2^1 \dots q_2^n$ akzeptierende Läufe von \mathcal{M}_1 und \mathcal{M}_2 .

Also ist $w \in L(\mathcal{M}_1)$ und $w \in L(\mathcal{M}_2)$.

„ $L(\mathcal{M}_1 \otimes \mathcal{M}_2) \supseteq L(\mathcal{M}_1) \cap L(\mathcal{M}_2)$ “: Analoge Schlussfolgerungen in entgegengesetzter Richtung. \square

Beispiel



Abschluss unter Komplement

Satz: Wenn L , L_1 und L_2 reguläre Sprachen sind, dann beschreiben auch die folgenden Ausdrücke reguläre Sprachen:

- (1) $L_1 \cup L_2$ (Abschluss unter Vereinigung)
- (2) $L_1 \cap L_2$ (Abschluss unter Schnitt)
- (3) \bar{L} **Abschluss unter Komplement**
- (4) $L_1 \circ L_2$ (Abschluss unter Konkatenation)
- (5) L^* (Abschluss unter Kleene-Stern)

Komplementoperator für DFAs

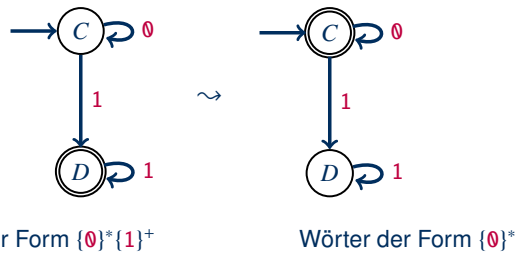
Idee: Wir können DFA komplementieren, indem wir akzeptierende und verwerfende Zustände vertauschen.

Für einen DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ sei $\overline{\mathcal{M}}$ der DFA $\langle Q, \Sigma, \delta, q_0, Q \setminus F \rangle$.

Beispiel:



Gegenbeispiel



Das Wort **010** zum Beispiel wird von keinem der beiden Automaten erkannt.

↪ keine komplementären Sprachen

Korrektheitsbeweis Komplementierung

Behauptung: Für jeden DFA gilt $L(\overline{\mathcal{M}}) = \overline{L(\mathcal{M})}$.

Beweis: „ $L(\overline{\mathcal{M}}) \subseteq \overline{L(\mathcal{M})}$ “: Sei $w = a_1 \dots a_n \in L(\overline{\mathcal{M}})$.

- Dann gibt es einen akzeptierenden Lauf $q_0 q_1 \dots q_n$ für w in $\overline{\mathcal{M}}$.
- Dann ist $q_n \in Q \setminus F$.
- Dann ist $q_0 q_1 \dots q_n$ ein verwerfender Lauf für w in \mathcal{M} .

Also ist $w \notin L(\mathcal{M})$, d.h. $w \in \overline{L(\mathcal{M})}$.

„ $L(\overline{\mathcal{M}}) \supseteq \overline{L(\mathcal{M})}$ “: Sei $w = a_1 \dots a_n \in \overline{L(\mathcal{M})}$.

- Dann hat \mathcal{M} einen verwerfenden Lauf $p_0 p_1 \dots p_m$ für w .
- Dann ist (1) $p_m \notin F$ oder (2) $m < |w|$.
- In Fall (1) ist $p_0 p_1 \dots p_m$ ein akzeptierender Lauf für w in $\overline{\mathcal{M}}$.
- In Fall (2) ... ?
... vielleicht gar kein akzeptierender Lauf für w in $\overline{\mathcal{M}}$...

Die Behauptung ist falsch!

Korrekte Komplementierung

Satz: Für jeden DFA \mathcal{M} mit totaler Übergangsfunktion gilt $L(\overline{\mathcal{M}}) = \overline{L(\mathcal{M})}$.

Beweis: Wie vorher für die falsche Behauptung, aber diesmal kann der Problemfall (2) der Rückrichtung nicht auftreten. □

Daraus folgt der Abschluss regulärer Sprachen unter Komplement, da wir jeden DFA mit totalen Übergängen ausstatten können.

Auch NFAs dürfen nicht direkt komplementiert werden:

- DFAs sind NFAs, unser Gegenbeispiel trifft weiterhin zu.
- Selbst NFAs, in denen jeder Zustand für jedes Symbol mindestens einen Folgezustand hat, können nicht „einfach so“ komplementiert werden. (Übung: Finden Sie ein Gegenbeispiel.)

Quiz: Komplementautomat

Quiz: Wir betrachten erneut den DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ und sein Komplement $\overline{\mathcal{M}}$, sowie den weiteren Automaten \mathcal{M}' :

...

Konkatenation von NFAs

Idee: Wir können Automaten „hintereinander hängen“, indem wir von Endzuständen des ersten zu Startzuständen des zweiten wechseln.

→ Besonders elegant geht das mit ϵ -Transitionen.

Gegeben seien zwei NFAs $\mathcal{M}_1 = \langle Q_1, \Sigma, \delta_1, Q_{0,1}, F_1 \rangle$ und $\mathcal{M}_2 = \langle Q_2, \Sigma, \delta_2, Q_{0,2}, F_2 \rangle$ mit $Q_1 \cap Q_2 = \emptyset$ (o.B.d.A.).

Der **Konkatenationsautomat** $\mathcal{M}_1 \circ \mathcal{M}_2$ ist ein ϵ -NFA gegeben durch $\langle Q_1 \cup Q_2, \Sigma, \delta, Q_{0,1}, F_2 \rangle$ mit:

$$\delta(q, a) = \begin{cases} \delta_1(q, a) & \text{falls } q \in Q_1 \\ \delta_2(q, a) & \text{falls } q \in Q_2 \end{cases} \quad \delta(q, \epsilon) = \begin{cases} Q_{0,2} & \text{falls } q \in F_1 \\ \emptyset & \text{andernfalls} \end{cases}$$

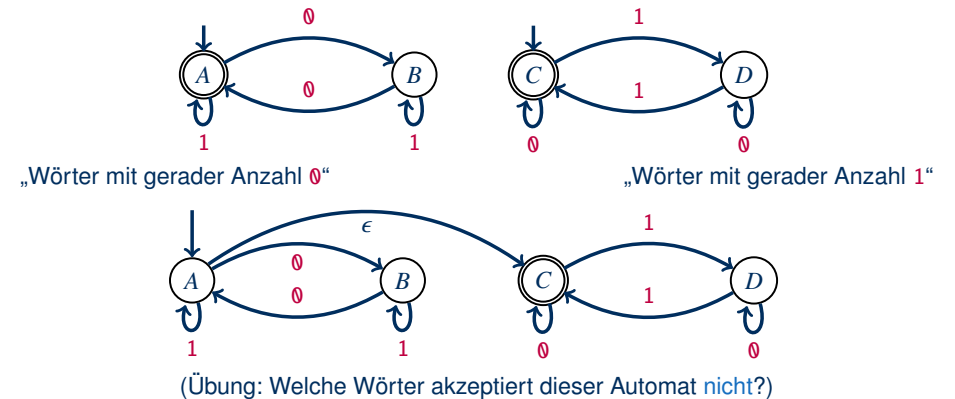
$\mathcal{M}_1 \circ \mathcal{M}_2$ simuliert also zunächst \mathcal{M}_1 . In jedem Endzustand aus F_1 entscheidet $\mathcal{M}_1 \circ \mathcal{M}_2$ nichtdeterministisch, diese Simulation fortzusetzen oder mit der Simulation von \mathcal{M}_2 zu beginnen.

Abschluss unter Konkatenation

Satz: Wenn L , L_1 und L_2 reguläre Sprachen sind, dann beschreiben auch die folgenden Ausdrücke reguläre Sprachen:

- (1) $L_1 \cup L_2$ (Abschluss unter Vereinigung)
- (2) $L_1 \cap L_2$ (Abschluss unter Schnitt)
- (3) \overline{L} (Abschluss unter Komplement)
- (4) $L_1 \circ L_2$ **Abschluss unter Konkatenation**
- (5) L^* (Abschluss unter Kleene-Stern)

Beispiel Konkatenation



Satz: Für alle NFA \mathcal{M}_1 und \mathcal{M}_2 gilt $\mathbf{L}(\mathcal{M}_1 \circ \mathcal{M}_2) = \mathbf{L}(\mathcal{M}_1) \circ \mathbf{L}(\mathcal{M}_2)$.

Der Beweis ist einfach und analog zu den bisher gezeigten.

Abschluss unter Kleene-Stern

Satz: Wenn L , L_1 und L_2 reguläre Sprachen sind, dann beschreiben auch die folgenden Ausdrücke reguläre Sprachen:

- (1) $L_1 \cup L_2$ (Abschluss unter Vereinigung)
- (2) $L_1 \cap L_2$ (Abschluss unter Schnitt)
- (3) \bar{L} (Abschluss unter Komplement)
- (4) $L_1 \circ L_2$ (Abschluss unter Konkatenation)
- (5) L^* **Abschluss unter Kleene-Stern**

Zusammenfassung und Ausblick

NFAs, DFAs, ϵ -NFAs und NFAs mit Wortübergängen beschreiben die selbe Klasse der regulären Sprachen.

Reguläre Sprachen sind **abgeschlossen unter** \cap , \cup , $\bar{}$, \circ , $*$ und allen davon ableitbaren Operatoren.

Den Sprachoperationen entsprechen **Operationen auf Automaten**. Manche erfordern bestimmte Typen von Automaten.

Offene Fragen:

- Gibt es noch mehr Darstellungsformen für reguläre Sprachen?
- Welche Sprachen sind nicht regulär?
- Wir haben gesehen, dass man Automaten manchmal vereinfachen kann – geht das noch besser?

Abschluss unter Kleene-Stern

Idee: Der Kleene-Stern ist eine verallgemeinerte Konkatenation, bei der ein Automat rekursiv hinter sich selbst „gehängt“ wird.

Gegeben sei ein NFA $M = \langle Q, \Sigma, \delta, Q_0, F \rangle$. Der Automat M^* ist ein ϵ -NFA gegeben durch $\langle Q', \Sigma, \delta', Q'_0, F' \rangle$ wobei

- $Q' = Q \cup \{q_\epsilon\}$ (wobei $q_\epsilon \notin Q$ o.B.d.A.)
- $\delta'(q, a) = \delta(q, a)$ für alle $q \in Q$, $a \in \Sigma$ und
 $\delta'(q_f, \epsilon) = Q_0$ für alle $q_f \in F$
- $Q'_0 = Q_0 \cup \{q_\epsilon\}$
- $F' = F \cup \{q_\epsilon\}$

Satz: $L(M^*) = L(M)^*$.

Der Beweis ist einfach und analog zu den bisher gezeigten.