Prof. J. Esparza
S. Sickert, J. Krämer


# Einführung in die theoretische Informatik

Sommersemester 2017 – Übungsblatt Lösungsskizze 4

**Übungsblatt**

Wir unterscheiden zwischen Übungs- und Abgabeblättern. Auf diesem *Übungsblatt* finden Sie eine Übersicht über die Kernaspekte, die Sie in Kalenderwoche 21 in den Tutorien diskutieren, üben und vertiefen. Die Aufgaben auf diesem Blatt dienen dem Üben und Verstehen des Vorlesungsstoffes, sowie dem *eigenständigen Erarbeiten* der Kernaspekte. Außerdem sollen Ihnen diese Aufgaben auch helfen, ein Gefühl dafür zu bekommen, was Sie inhaltlich in der Klausur erwartet. Klausuraufgaben können jedoch deutlich von den hier gestellten Aufgaben abweichen. Abschreiben und Auswendiglernen von Lösungen wird Ihnen daher keinen dauerhaften Erfolg in der Vorlesung bringen. Fragen zu den Übungsblättern können Sie montags bis donnerstags von 12 Uhr bis 14 Uhr in der *THEO-Sprechstunde* in Raum 03.11.034 stellen.

**Kernaspekte**

K4.1 korrektes Wiedergeben der folgenden Definitionen

- Wortproblem
- Leerheitsproblem
- Endlichkeitsproblem
- Äquivalenzproblem

K4.2 korrektes Wiedergeben des Pumping Lemma
K4.3 korrektes Wiedergeben des Ardens Lemma
K4.4 mithilfe des Pumping Lemmas beweisen, dass eine gegebene Sprache nicht regulär ist
K4.5 beweisen, dass eine gegebene Sprache regulär ist
K4.6 mithilfe von Homomorphismen auf Sprachen beweisen, dass eine gegebene Sprache (nicht) regulär ist.
K4.7 das Wortproblem/Leerheitsproblem/Endlichkeitsproblem/Äquivalenzproblem für einen gegebenen NFA/DFA mithilfe des entsprechenden Algorithmus aus der Vorlesung entscheiden
K4.8 das Wortproblem/Leerheitsproblem/Endlichkeitsproblem/Äquivalenzproblem für andere Kodierungen regulärer Sprachen durch passende Reduktion auf NFAs/DFAs entscheiden
K4.9 beurteilen, welche Kodierung zur Entscheidung des Wortproblem/Leerheitsproblem/Endlichkeitsproblem/Äquivalenzproblem am effizientesten ist
K4.10 zu einem gegebenen NFA N einen regulären Ausdruck r mithilfe eines Gleichungssystems (Ardens Lemma) angeben, so dass $\mathsf{L}(\mathsf{N}) = \mathsf{L}(\mathsf{r})$
K4.11 begründet entscheiden, ob gegebene Beispiele neu eingeführte Definitionen erfüllen
K4.12 Aussagen, mit neu eingeführten Definitionen beweisen oder widerlegen

**AUFGABE 4.1.** (*Produktkonstruktion*) B+C

Sei $\Sigma$ ein Alphabet und $\mathsf{A}, \mathsf{B}, \mathsf{C} \subseteq \Sigma^*$. Wir definieren den $\mathtt{ITE}$-Operator (*if-then-else*) wie folgt:

$$\mathtt{ITE}(\mathsf{A}, \mathsf{B}, \mathsf{C}) = \{w \in \Sigma^* \mid (w \in \mathsf{A} \wedge w \in \mathsf{B}) \vee (w \notin \mathsf{A} \wedge w \in \mathsf{C})\}$$

Seien jetzt $\mathsf{A}, \mathsf{B}, \mathsf{C} \subseteq \Sigma^*$ reguläre Sprachen, die von den DFAs $\mathsf{D}_\mathsf{A}$, $\mathsf{D}_\mathsf{B}$ und $\mathsf{D}_\mathsf{C}$ erkannt werden.

(a) Erweitern Sie die Produktkonstruktion aus der Vorlesung, so dass aus den Automaten $\mathsf{D}_\mathsf{A}$, $\mathsf{D}_\mathsf{B}$ und $\mathsf{D}_\mathsf{C}$ ein DFA $\mathsf{D}_{\mathtt{ITE}(\mathsf{A},\mathsf{B},\mathsf{C})}$ mit $\mathsf{L}(\mathsf{D}_{\mathtt{ITE}(\mathsf{A},\mathsf{B},\mathsf{C})}) = \mathtt{ITE}(\mathsf{A}, \mathsf{B}, \mathsf{C})$ konstruiert wird und beweisen Sie die Korrektheit Ihrer Konstruktion!

(b) Sei nun $\Sigma = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ und $\mathsf{A} = \{w \in \Sigma^* \mid |w| \text{ ist gerade}\}$, $\mathsf{B} = \{\mathsf{a}^n\mathsf{b}^m \mid n, m \geq 0\}$ und $\mathsf{C} = \{\mathsf{c}^n\mathsf{b}^m \mid n, m \geq 0\}$. Geben Sie einen DFA D für $\mathtt{ITE}(\mathsf{A}, \mathsf{B}, \mathsf{C})$, indem Sie Ihre Konstruktion aus Aufgabenteil (a) verwenden.

**Hinweis:** Verwenden Sie $\mathsf{Q}_{\mathtt{ITE}(\mathsf{A},\mathsf{B},\mathsf{C})} = \mathsf{Q}_\mathsf{A} \times \mathsf{Q}_\mathsf{B} \times \mathsf{Q}_\mathsf{C}$ als Menge aller Zustände für die erweiterte Produktkonstruktion in Aufgabenteil (a).

*Lösungsskizze*

(a) **Konstruktion**

$$\begin{aligned}
\mathsf{D}_{\mathtt{ITE}(\mathsf{A},\mathsf{B},\mathsf{C})} &\coloneqq (\mathsf{Q}^{\mathsf{A}} \times \mathsf{Q}^{\mathsf{B}} \times \mathsf{Q}^{C}, \Sigma, \delta_\times, \mathsf{q}_{0,\times}, \mathsf{F}_\times) \\
\delta_\times((q, q', q''), \mathsf{a}) &\coloneqq (\delta^{\mathsf{A}}(q, \mathsf{a}), \delta^{\mathsf{B}}(q', \mathsf{a}), \delta^{\mathsf{C}}(q'', \mathsf{a})) \\
\mathsf{q}_{0,\times} &\coloneqq (\mathsf{q}_0^{\mathsf{A}}, \mathsf{q}_0^{\mathsf{B}}, \mathsf{q}_0^{\mathsf{C}}) \\
\mathsf{F}_\times &\coloneqq \{(q^{\mathsf{A}}, q^{\mathsf{B}}, q^{\mathsf{C}}) \mid (q^{\mathsf{A}} \in \mathsf{F}^{\mathsf{A}} \wedge q^{\mathsf{B}} \in \mathsf{F}^{\mathsf{B}}) \vee (q^{\mathsf{A}} \notin \mathsf{F}^{\mathsf{A}} \wedge q^{\mathsf{C}} \in \mathsf{F}^{\mathsf{C}})\}
\end{aligned}$$

**Korrektheit** Mit Induktion über $w \in \Sigma^*$ erhalten wir für alle $q, q', q''$:

$$\hat{\delta_\times}((q, q', q''), w) = (\hat{\delta^{\mathsf{A}}}(q, w), \hat{\delta^{\mathsf{B}}}(q', w), \hat{\delta^{\mathsf{C}}}(q'', w))$$

Damit gilt:

$$\begin{aligned}
& w \in \mathsf{L}(\mathsf{D}_{\mathtt{ITE}(\mathsf{A},\mathsf{B},\mathsf{C})}) \\
\Leftrightarrow\ & \hat{\delta_\times}(\mathsf{q}_{0,\times}, w) \in \mathsf{F}_\times \\
\Leftrightarrow\ & (\hat{\delta^{\mathsf{A}}}(\mathsf{q}_0^{\mathsf{A}}, w) \in \mathsf{F}^{\mathsf{A}} \wedge \hat{\delta^{\mathsf{B}}}(\mathsf{q}_0^{\mathsf{B}}, w) \in \mathsf{F}^{\mathsf{B}}) \vee (\hat{\delta^{\mathsf{A}}}(\mathsf{q}_0^{\mathsf{A}}, w) \notin \mathsf{F}^{\mathsf{A}} \wedge \hat{\delta^{\mathsf{C}}}(\mathsf{q}_0^{\mathsf{C}}, w) \in \mathsf{F}^{\mathsf{C}}) \\
\Leftrightarrow\ & (w \in \mathsf{A} \wedge w \in \mathsf{B}) \vee (w \notin \mathsf{A} \wedge w \in \mathsf{C}) \\
\Leftrightarrow\ & w \in \mathtt{ITE}(\mathsf{A}, \mathsf{B}, \mathsf{C})
\end{aligned}$$

(b) Siehe Abbildung 1. Alle nicht-gezeichneten Kanten, die für eine totale Transitionsrelation notwendig sind, führen zu einem Fehlerzustand.

**AUFGABE 4.2.** (*Pumping Lemma*) C

Beweisen Sie für jede der folgenden Sprachen mithilfe des Pumping Lemmas, dass sie *nicht* regulär sind.

(a) $\mathsf{L}_1 = \{w \in \{0,1\}^* \mid w = w^R\}$
(b) $\mathsf{L}_2 = \{w \in \{0,1\}^* \mid |w|_0 \geq |w|_1\}$
(c) $\mathsf{L}_3 = \{\varepsilon, \mathsf{a}, \mathsf{a}^{n \cdot m} \mid m \in \mathsf{N}, n \in \mathsf{N}, m > 1, n > 1\}$
(d) $\mathsf{L}_4 = \{\mathsf{a}^{6i}\mathsf{b}^{6i} \mid i \geq 0\}$
(e) $\mathsf{L}_5 = \{\mathsf{a}^{2^i} \mid i \geq 0\}$

*Lösungsskizze*

(a) Angenommen, $\mathsf{L}_1$ wäre regulär. Dann können wir das Pumping-Lemma anwenden.
Sei $n \in \mathbb{N}$ eine Pumping-Lemma-Zahl. Dann gilt $0^n10^n \in \mathsf{L}_1$. Es gibt also für $z = 0^n10^n$ eine Zerlegung $z = uvw$ mit $v \neq \varepsilon$ und $|uv| \leq n$, so dass $(*)$ $uv^iw \in \mathsf{L}_1$ für alle $i \in \mathbb{N}_0$. Deshalb $uv = 0^k$ für $0 < k \leq n$. Also $u = 0^i$ und $v = 0^j$ mit $i + j = k$ und $j > 0$.
Dann ist aber $uv^2w = 0^i0^{2j}0^{n-k}10^n \notin \mathsf{L}_1$, denn $i + 2j + n - k > n$. Dies steht im Widerspruch zu $(*)$, d.h. unsere ursprüngliche Annahme, dass $\mathsf{L}_1$ regulär ist, ist falsch.

(b) Angenommen, $\mathsf{L}_2$ wäre regulär. Dann können wir das Pumping-Lemma anwenden.
Sei also $n \in \mathbb{N}$ eine Pumping-Lemma-Zahl. Dann gilt $z = 0^n1^n \in \mathsf{L}_2$. Es gibt daher eine Zerlegung $z = uvw$ mit $v \neq \varepsilon$, $|uv| \leq n$ und $uv^iw \in \mathsf{L}_2$ für alle $i \in \mathbb{N}_0$. Aus unserer Wahl von $z$ folgt, dass $u = 0^i$ und $v[j]$ mit $j > 0$ gilt.
Allerdings ist $uv^0w = 0^i0^{n-i-j}1^n \notin \mathsf{L}_2$, ein Widerspruch zum Pumping-Lemma. $\mathsf{L}_2$ ist somit nicht regulär.

(c) Angenommen, $\mathsf{L}_3$ wäre regulär. Da die regulären Sprachen unter Komplement abgeschlossen sind, ist auch

$$\overline{\mathsf{L}_3} = \{\mathsf{a}^n \mid n \in \mathbb{N}, n > 1, n \text{ ist prim}\} =: \mathsf{L}$$

regulär. Dann können wir das Pumping-Lemma anwenden.
Sei $n$ eine Pumping-Lemma-Zahl. Da die Menge der Primzahlen bekanntlich unendlich ist, gibt es eine Primzahl $p$ bzw. ein $z = 1^p \in \mathsf{L}_3$, so dass $z$ mindestens $n$ Zeichen enthält, also $|z| \geq n$. Sei also $z \in \mathsf{L}$ mit $|z| \geq n$ und sei $z = uvw$ mit $|uv| \leq n$, $v$ nicht leer und $uv^iw \in \mathsf{L}$ für alle $i \geq 0$. Wegen $uw = uv^0w \in \mathsf{L}$ ist $|uw|$ eine Primzahl. Wir setzen $\mathsf{x} = uv^{|uw|}w$.
Es gilt $\mathsf{x} \in \mathsf{L}$, d. h. $|\mathsf{x}|$ ist eine Primzahl. Es folgt aber auch $|\mathsf{x}| = |uw| + |uw|\,|v| = |uw|\,(1 + |v|)$, d.h. $|\mathsf{x}|$ ist wegen $|v| \neq 0$ keine Primzahl. Widerspruch. $\mathsf{L}_3$ ist somit nicht regulär.

(d) Angenommen: $\mathsf{L}_4$ wäre regulär. Dann können wir das Pumping-Lemma anwenden.
Sei $n \geq 1$ eine Pumping-Lemma-Zahl. Dann gilt $\mathsf{a}^{6n}\mathsf{b}^{6n} \in \mathsf{L}_4$ und $\left|\mathsf{a}^{6n}\mathsf{b}^{6n}\right| \geq n$. Es gibt also für $z = \mathsf{a}^{6n}\mathsf{b}^{6n}$ eine Zerlegung $z = uvw$ mit $v \neq \varepsilon$ und $|uv| \leq n$, so dass $(*)$ $uv^iw \in \mathsf{L}_4$ für alle $i \in \mathbb{N}_0$. Deshalb $v = \mathsf{a}^j$ für ein $0 < j \leq n$.
Dann ist aber $uv^0w = \mathsf{a}^{6n-j}\mathsf{b}^{6n} \notin \mathsf{L}_4$, denn $6n - j < 6n$. Dies steht im Widerspruch zu $(*)$, d.h. unsere ursprüngliche Annahme, dass $\mathsf{L}_4$ regulär ist, ist falsch.

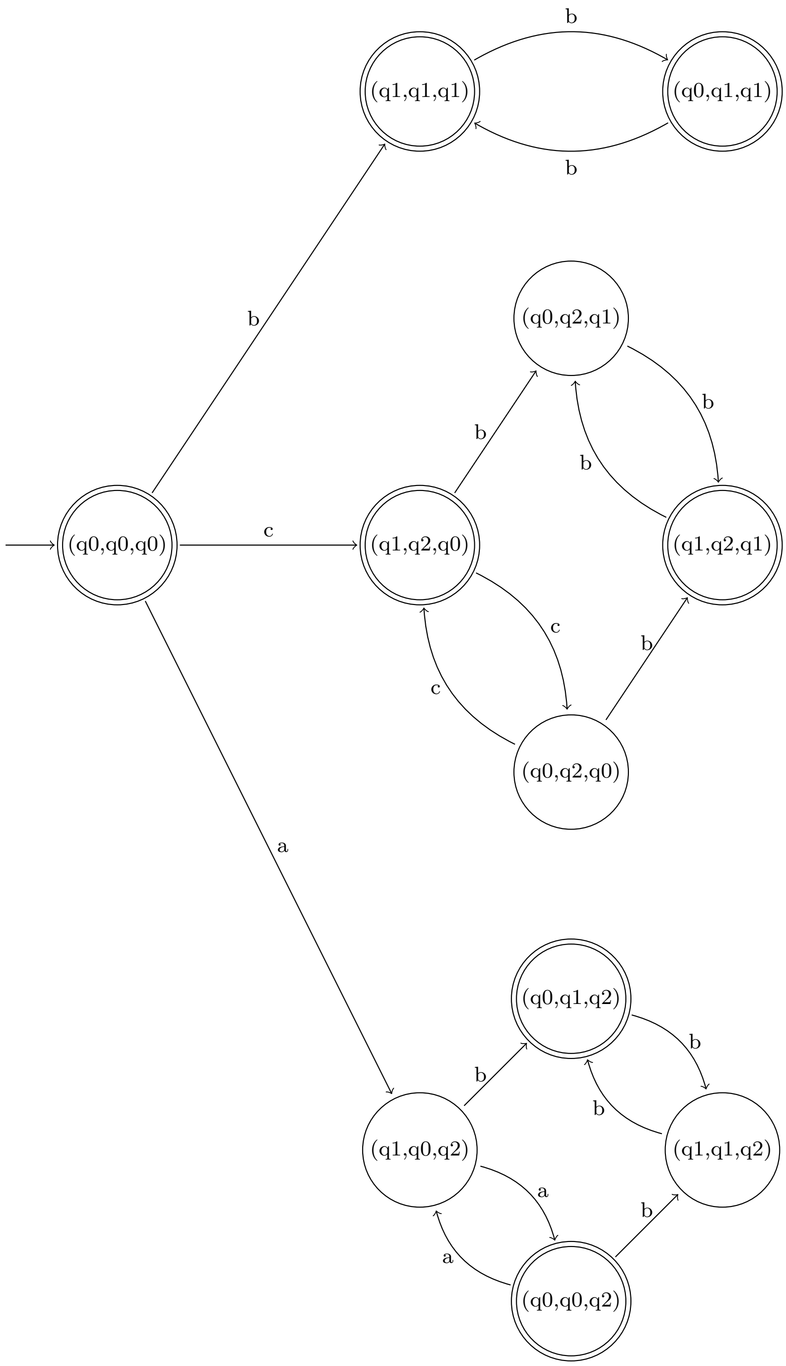



**Abbildung 1** – Automat zu Aufgabe 4.1b)

(e) Angenommen: $\mathsf{L}_5$ wäre regulär. Dann können wir das Pumping-Lemma anwenden.
Sei $n \geq 1$ eine Pumping-Lemma-Zahl. Dann gilt $\mathsf{a}^{2^n} \in \mathsf{L}_5$ und $\left|\mathsf{a}^{2^n}\right| \geq n$. Es gibt also für $z = \mathsf{a}^{2^n}$ eine Zerlegung $z = uvw$ mit $v \neq \varepsilon$ und $|uv| \leq n$, so dass $(*)$ $uv^iw \in \mathsf{L}_5$ für alle $i \in \mathbb{N}_0$. Deshalb $v = \mathsf{a}^j$ für ein $0 < j \leq n$.
Dann ist aber $uv^2w = \mathsf{a}^{2^n+j} \notin \mathsf{L}_5$, denn

$$2^n < 2^n + j \leq 2^n + n < 2^n + 2^n = 2^{n+1}$$

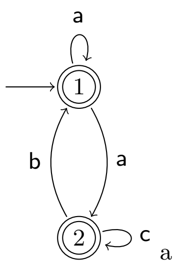
Dies steht im Widerspruch zu $(*)$, d.h. unsere ursprüngliche Annahme, dass $\mathsf{L}_5$ regulär ist, ist falsch.

**AUFGABE 4.3.** (*Ardens Lemma*) C

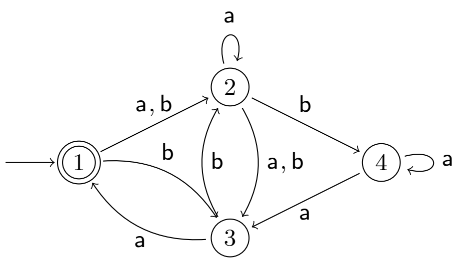
Betrachten Sie die beiden NFAs $\mathsf{N}_a$ und $\mathsf{N}_b$.

(a) Geben Sie für beide NFAs das nach Vorlesung dazugehörige Gleichungssystem an.
(b) Geben Sie dann für jeden der beiden NFAs mittels Gauß-Verfahren und Ardens Lemma einen regulären

Ausdruck $r_a$ bzw. $r_b$ an, so dass $L(N_a) = L(r_a)$ bzw. $L(N_b) = L(r_b)$




**(a)** NFA $N_a$




**(b)** NFA $N_b$

*Lösungsskizze*

Die Variable $X_i$ für einen Zustand $i$ wird gleichgesetzt mit der Vereinigung der Variablen der in einem Schritt erreichbaren Zuständen mit der Kantenbeschriftung als Präfix. Zusätzlich fügen wir die Konstante $\{\varepsilon\}$ für Endzustände ein.

(a) Aufstellen des Gleichungssystem:

$$\begin{aligned} X_1 &\equiv \varepsilon \mid aX_1 \mid aX_2 \\ X_2 &\equiv \varepsilon \mid bX_1 \mid cX_2 \end{aligned}$$

Lösen nach $X_1$:

$$\begin{aligned} X_2 &\equiv cX_2 \mid (\varepsilon \mid bX_1) && \text{Umformen} \\ X_2 &\equiv c^*(\varepsilon \mid bX_1) && \text{Ardens Lemma} \\ X_1 &\equiv \varepsilon \mid aX_1 \mid a(c^*(\varepsilon \mid bX_1)) && \text{Einsetzen von } X_2 \\ X_1 &\equiv (a \mid ac^*b)X_1 \mid (\varepsilon \mid ac^*) && \text{Umformen} \\ X_1 &\equiv (a \mid ac^*b)^*(\varepsilon \mid ac^*) && \text{Ardens Lemma} \end{aligned}$$

(b) Aufstellen des Gleichungssystem:

$$\begin{aligned} X_1 &\equiv \varepsilon \mid (a|b)X_2 \mid bX_3 \\ X_2 &\equiv aX_2 \mid (a|b)X_3 \mid bX_4 \\ X_3 &\equiv aX_1 \mid bX_2 \\ X_4 &\equiv aX_3 \mid aX_4 \end{aligned}$$

Lösen nach $X_1$:

$$\begin{aligned} X_1 &\equiv \varepsilon \mid baX_1 \mid (a|b|bb)X_2 && \text{Einsetzen von } X_3 \\ X_2 &\equiv (aa|ba)X_1 \mid (a|ab|bb)X_2 \mid bX_4 && \text{Einsetzen von } X_3 \\ X_4 &\equiv aaX_1 \mid abX_2 \mid aX_4 && \text{Einsetzen von } X_3 \\ X_4 &\equiv a^*(aaX_1 \mid abX_2) && \text{Ardens Lemma} \\ X_2 &\equiv (aa|baa^*)X_1 \mid (a|ab|ba^*b)X_2 && \text{Einsetzen von } X_4 \\ X_2 &\equiv (a|ab|ba^*b)^*(aa|baa^*)X_1 && \text{Ardens Lemma} \\ X_1 &\equiv \varepsilon \mid (ba \mid (a|b|bb)(a|ab|ba^*b)^*(aa|baa^*))X_1 && \text{Einsetzen von } X_2 \\ X_1 &\equiv (ba \mid (a|b|bb)(a|ab|ba^*b)^*(aa|baa^*))^* && \text{Ardens Lemma} \end{aligned}$$

**AUFGABE 4.4.** (*Umkehrung(Spiegelung)*) D

Ziel dieser Aufgabe ist es Algorithmen anzugeben, die es erlauben, die Umkehrung einer Sprache zu berechnen, d.h. für jede reguläre Sprache $L$ die Sprache $L^R$ anzugeben, so dass $L^R := \{w^R \mid w \in L\}$.

(a) Sei D ein DFA. Geben Sie einen Algorithmus an, der den DFA D in einen $\varepsilon$-NFA N übersetzt, so dass $L(D)^R = L(N)$. Beweisen Sie auch die Korrektheit Ihres Verfahrens.

(b) Konstruieren Sie für den regulären Ausdruck $r = 0 \mid 1(0|1)^*0$ einen DFA und wenden Sie Ihr Spiegelungsverfahren aus Aufgabenteil (a) an.

(c) Geben Sie einen regulären Ausdruck $r'$ mit $L(r) = L(r')^R$ an.

(d) Geben Sie nun ein Verfahren an, welches rekursiv einen regulären Ausdruck $r$ in einen regulären Ausdruck $r'$ umschreibt, so dass $L(r)^R = L(r')$ gilt. Zeigen Sie die Korrektheit Ihres Verfahrens mittels struktureller Induktion und seien Sie besonders ausführlich in den Fällen Konkatenation, Vereinigung und Stern.

*Lösungsskizze*

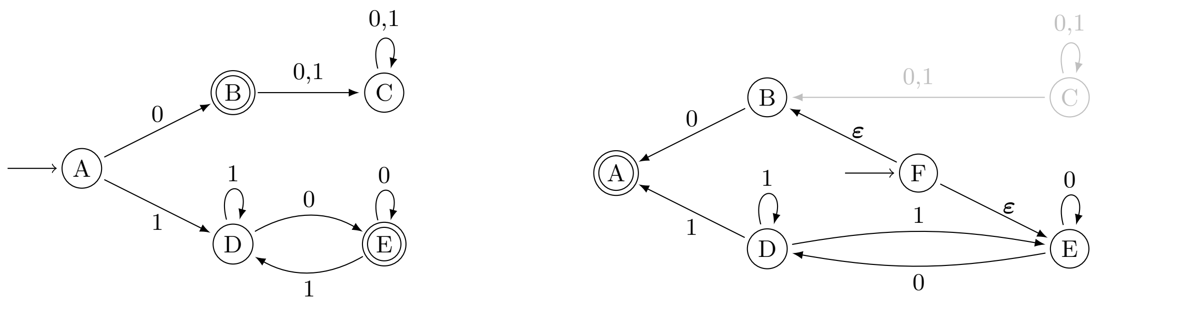
(a) Sei $\mathsf{D} = (\mathsf{Q}, \Sigma, \delta, \mathsf{q}_0, \mathsf{F})$ ein DFA.

**Konstruktion:** Wir definieren den $\varepsilon$-NFA $\mathsf{N}' = (\mathsf{Q}', \Sigma', \delta', \mathsf{q}_0', \mathsf{F}')$ wie folgt:

- $\mathsf{Q}' \coloneqq \mathsf{Q} \cup \{\hat{q}\}$, wobei $\hat{q} \notin \mathsf{Q}$
- $\mathsf{q}_0' \coloneqq \hat{q}$
- $\Sigma' \coloneqq \Sigma \cup$
- $(\hat{q}, \varepsilon, q') \in \delta'$ für alle $q' \in \mathsf{F}$
- $(q', \mathsf{x}, q) \in \delta'$ für $\mathsf{x} \in \Sigma$ und $(q, \mathsf{x}, q') \in \delta$
- $\mathsf{F} = \{\mathsf{q}_0\}$

**Korrektheit:** Wir müssen zeigen, dass $w \in \mathsf{L}(\mathsf{D})^R$. Nach Definition der Umkehrsprache und da $(w^R)^R = w$ und $(\mathsf{L}^R)^R = \mathsf{L}$, wissen wir $w \in \mathsf{L}(\mathsf{D})^R$ genau dann, wenn $w^R \in \mathsf{L}(\mathsf{D})$. Wir können dann mit Induktion über die Länge von $w$ beweisen, dass $w \in \mathsf{L}(\mathsf{D}) \Leftrightarrow w^R \in \mathsf{L}(\mathsf{N})$.

(b) Siehe Abbildung:




**(a)** DFA für $\mathsf{L}(\mathsf{r})$

**(b)** $\varepsilon$-NFA für $\mathsf{L}(\mathsf{r})^R$, in Grau die nicht-erreichbaren Teile des Automaten

(c) $\mathsf{r}' = 0 \mid 0(0|1)^*1$

(d) **Konstruktion:** Wir definieren folgende rekursive Prozedur:

$$\begin{aligned} &rev(\emptyset) = \emptyset && rev(\alpha\beta) = rev(\beta)rev(\alpha) \\ &rev(\varepsilon) = \varepsilon && rev(\alpha \mid \beta) = rev(\alpha) \mid rev(\beta) \\ &rev(\mathsf{x}) = \mathsf{x} \quad \text{für } \mathsf{x} \in \Sigma && rev(\alpha^*) = rev(\alpha)^* \end{aligned}$$

**Korrektheit:** Da für alle $u, v \in \Sigma^*$ gilt: $(uv)^R = (v_m \dots v_1 u_n \dots u_1) = v^R u^R$, folgt für beliebige Sprachen $\mathsf{A}$ und $\mathsf{B}$: $(\mathsf{AB})^R = \mathsf{B}^R\mathsf{A}^R$ und $(\mathsf{A}^R)^* = (\mathsf{A}^*)^R$. Dann können wir die Korrektheit, d.h. $\mathsf{L}(rev(\gamma)) = \mathsf{L}(\gamma)^R$ mit struktureller Induktion über $\gamma$ zeigen:

$\gamma = \emptyset$: $\mathsf{L}(rev(\emptyset)) = \emptyset = \emptyset^R = \mathsf{L}(\emptyset)^R$

$\gamma = \varepsilon$: $\mathsf{L}(rev(\varepsilon)) = \{\varepsilon\} = \{\varepsilon\}^R = \mathsf{L}(\varepsilon)^R$

$\gamma = \mathsf{x}$: $\mathsf{L}(rev(\mathsf{x})) = \{\mathsf{x}\} = \{\mathsf{x}\}^R = \mathsf{L}(\mathsf{x})^R$ für $\mathsf{x} \in \Sigma$

$\gamma = \alpha\beta$: $\mathsf{L}(rev(\alpha\beta)) = \mathsf{L}(rev(\beta))\mathsf{L}(rev(\alpha)) \overset{\mathsf{IH}}{=} \mathsf{L}(\beta)^R\mathsf{L}(\alpha)^R = \mathsf{L}(\alpha\beta)^R$

$\gamma = \alpha \mid \beta$: $\mathsf{L}(rev(\alpha \mid \beta)) = \mathsf{L}(rev(\alpha)) \cup \mathsf{L}(rev(\beta)) \overset{\mathsf{IH}}{=} \mathsf{L}(\alpha)^R \cup \mathsf{L}(\beta)^R = \mathsf{L}(\alpha \mid \beta)^R$

$\gamma = \alpha^*$: $\mathsf{L}(rev(\alpha^*)) = \mathsf{L}(rev(\alpha))^* \overset{\mathsf{IH}}{=} (\mathsf{L}(\alpha)^R)^* = \mathsf{L}(\alpha^*)^R$

**AUFGABE 4.5.** (*Leere Sprachen*) D

Sie haben in der Vorlesung gesehen, wie man entscheiden kann, ob ein endlicher Automat die leere Sprache akzeptiert. Wir entwickeln jetzt ein Verfahren, das direkt und ohne Umwege über endliche Automaten prüft, ob ein regulärer Ausdruck die leere Sprache beschreibt.

Definieren Sie hierzu eine rekursive Funktion $iszero : \mathsf{RE} \mapsto \mathbb{B}$ über den regulären Ausdrücken und beweisen Sie, dass für alle Ausdrücke $\mathsf{r}$ gilt:

$$iszero(\mathsf{r}) \Leftrightarrow \mathsf{L}(\mathsf{r}) = \emptyset$$

*Lösungsskizze*

**Konstruktion:** Wir definieren folgende rekursive Prozedur:

$$\begin{aligned}
&\mathit{iszero}(\boldsymbol{\emptyset}) = \mathsf{true} && && \mathit{iszero}(\alpha\beta) = \mathit{iszero}(\alpha) \vee \mathit{iszero}(\beta)\\
&\mathit{iszero}(\boldsymbol{\varepsilon}) = \mathsf{false} && && \mathit{iszero}(\alpha \mid \beta) = \mathit{iszero}(\alpha) \wedge \mathit{iszero}(\beta)\\
&\mathit{iszero}(\mathsf{x}) = \mathsf{false} && \text{für } \mathsf{x} \in \Sigma && \mathit{iszero}(\alpha^*) = \mathsf{false}
\end{aligned}$$

**Korrektheit** Wir zeigen $\mathit{iszero}(\gamma) \Leftrightarrow \mathsf{L}(\gamma) = \emptyset$ mit struktureller Induktion über $\gamma$:

$\gamma = \boldsymbol{\emptyset}$: $\mathit{iszero}(\boldsymbol{\emptyset}) \Leftrightarrow \mathsf{true} \Leftrightarrow \mathsf{L}(\boldsymbol{\emptyset}) = \emptyset$

$\gamma = \boldsymbol{\varepsilon}$: $\mathit{iszero}(\boldsymbol{\varepsilon}) \Leftrightarrow \mathsf{false} \Leftrightarrow \mathsf{L}(\boldsymbol{\varepsilon}) = \{\varepsilon\} = \emptyset$

$\gamma = \mathsf{x}$: $\mathit{iszero}(\mathsf{x}) \Leftrightarrow \mathsf{false} \Leftrightarrow \mathsf{L}(\mathsf{x}) = \{\mathsf{x}\} = \emptyset$ für $\mathsf{x} \in \Sigma$

$\gamma = \alpha\beta$: $\mathit{iszero}(\alpha\beta) \Leftrightarrow \mathit{iszero}(\alpha) \vee \mathit{iszero}(\beta) \overset{\mathsf{IH}}{\Leftrightarrow} \mathsf{L}(\alpha) = \emptyset \vee \mathsf{L}(\beta) = \emptyset \Leftrightarrow \mathsf{L}(\alpha\beta) = \emptyset$.

$\gamma = \alpha \mid \beta$: $\mathit{iszero}(\alpha \mid \beta) \Leftrightarrow \mathit{iszero}(\alpha) \wedge \mathit{iszero}(\beta) \overset{\mathsf{IH}}{\Leftrightarrow} \mathsf{L}(\alpha) = \emptyset \wedge \mathsf{L}(\beta) = \emptyset \Leftrightarrow \mathsf{L}(\alpha|\beta) = \emptyset$.

$\gamma = \alpha^*$: $\mathit{iszero}(\alpha^*) \Leftrightarrow \mathsf{false} \Leftrightarrow \mathsf{L}(\alpha^*) = \emptyset$

**AUFGABE 4.6.** (*Homomorphismen auf regulären Sprachen*) E / C

Sei $\Sigma$ ein Alphabet und $\mathsf{L} \subseteq \Sigma^*$ eine reguläre Sprache. Weiter sei $h\colon \Sigma \to \Sigma^*$ eine Abbildung, die jedem Zeichen $\mathsf{x} \in \Sigma$ ein Wort $h(\mathsf{x}) \in \Sigma^*$ zuordnet. $h$ erweitert man kanonisch auf $\Sigma^*$ mittels $h(\varepsilon) = \varepsilon$ und $h(wv) = h(w)h(v)$ für alle $w, v \in \Sigma^*$. Schließlich sei $h(\mathsf{L}) = \{h(w) \mid w \in \mathsf{L}\}$.

(a) Zeigen Sie: Ist $\mathsf{L}$ regulär, dann ist auch $h(\mathsf{L})$ regulär.

(b) Zeigen Sie unter Verwendung des Resultats aus Aufgabenteil (a), dass $\mathsf{L}' = \{\mathsf{ab}^{3i}\mathsf{cd}^{2i}\mathsf{e} \mid i \in \mathbb{N}_0\}$ nicht regulär ist.

*Lösungsskizze*

(a) **Konstruktion:** Wir liften $h$ auf reguläre Ausdrücke und erhalten $h_{RE}$:

$$\begin{aligned}
&h_{RE}(\emptyset) = \emptyset && && h_{RE}(\alpha\beta) = h_{RE}(\alpha)h_{RE}(\beta)\\
&h_{RE}(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon} && && h_{RE}(\alpha \mid \beta) = h_{RE}(\alpha) \mid h_{RE}(\beta)\\
&h_{RE}(\mathsf{x}) = h(\mathsf{x}) && \text{für } \mathsf{x} \in \Sigma && h_{RE}(\alpha^*) = h_{RE}(\alpha)^*
\end{aligned}$$

**Korrektheit:** Wir zeigen $\mathsf{L}(h_{RE}(\gamma)) = h(\mathsf{L}(\gamma))$ mit struktureller Induktion über $\gamma$. Dabei verwenden wir das $h$ strukturerhaltend ist: $h(\mathsf{A}^n) = (h(\mathsf{A}))^n$, sowie $h(\mathsf{A} \cup \mathsf{B}) = h(\mathsf{A}) \cup h(\mathsf{B})$.

$\gamma = \emptyset$: $\mathsf{L}(h_{RE}(\emptyset)) = \emptyset = h(\emptyset)$

$\gamma = \varepsilon$: $\mathsf{L}(h_{RE}(\varepsilon)) = \{\varepsilon\} = h(\{\varepsilon\})$

$\gamma = \mathsf{x}$: $\mathsf{L}(h_{RE}(\mathsf{x})) = \{h(\mathsf{x})\} = h(\{\mathsf{x}\})$ für $\mathsf{x} \in \Sigma$.

$\gamma = \alpha\beta$: $\mathsf{L}(h_{RE}(\alpha\beta)) = \mathsf{L}(h_{RE}(\alpha))\mathsf{L}(h_{RE}(\beta)) \overset{\mathsf{IH}}{=} h(\mathsf{L}(\alpha))h(\mathsf{L}(\beta)) = h(\mathsf{L}(\alpha)\mathsf{L}(\beta)) = h(\mathsf{L}(\alpha\beta))$

$\gamma = \alpha \mid \beta$: $\mathsf{L}(h_{RE}(\alpha \mid \beta)) = \mathsf{L}(h_{RE}(\alpha)) \cup \mathsf{L}(h_{RE}(\beta)) \overset{\mathsf{IH}}{=} h(\mathsf{L}(\alpha)) \cup h(\mathsf{L}(\beta)) = h(\mathsf{L}(\alpha) \cup \mathsf{L}(\beta)) = h(\mathsf{L}(\alpha \mid \beta))$

$\gamma = \alpha^*$: $\mathsf{L}(h_{RE}(\alpha^*)) = \mathsf{L}(h_{RE}(\alpha))^* \overset{\mathsf{IH}}{=} h(\mathsf{L}(\alpha))^* = \bigcup_{i \geq 0} (h(\mathsf{L}(\alpha)))^i = \bigcup_{i \geq 0} h(\mathsf{L}(\alpha)^i) = h(\mathsf{L}(\alpha^*))$

Sei $\mathsf{r}$ ein regulärer Ausdruck für $\mathsf{L}$. Dann

$$h(\mathsf{L}) = h(\mathsf{L}(\mathsf{r})) = \mathsf{L}(h_{RE}(\mathsf{r}))$$

und somit $h(\mathsf{L})$ regulär.

*Alternativ kann der Beweis anhand eines Automaten geführt werden. Dabei werden die Kantenbeschriftungen mit h umgeschrieben.*

(b) Wir definieren eine Abbildung $h\colon \Sigma \to \Sigma^*$:

$$h(\mathsf{a}) = \varepsilon \qquad h(\mathsf{b}) = \mathsf{aa} \qquad h(\mathsf{c}) = \varepsilon \qquad h(\mathsf{d}) = \mathsf{bbb} \qquad h(\mathsf{e}) = \varepsilon.$$

Wir nehmen an das $\mathsf{L}'$ regulär ist. Dann ist laut (a) auch

$$h(\{\mathsf{ab}^{3i}\mathsf{cd}^{2i}\mathsf{e} \mid i \in \mathbb{N}_0\}) = \{\mathsf{a}^{6i}\mathsf{b}^{6i} \mid i \in \mathbb{N}_0\}$$

regulär. Widerspruch zu Aufgabe 4.2 d). $\mathsf{L}'$ ist somit nicht regulär.