TECHNISCHE UNIVERSITÄT DRESDEN

FORMALE SYSTEME

ÜBUNG 4

Eric Kunze
eric.kunze@tu-dresden.de

TU Dresden, 11. November 2021

REGULÄRE AUSDRÜCKE

Die Menge der **regulärer Ausdrücke** über einem Alphabet Σ ist induktiv wie folgt definiert:

- $\emptyset$ ist ein regulärer Ausdruck
- ε ist ein regulärer Ausdruck
- $\mathtt{a}$ ist ein regulärer Ausdruck für jedes $\mathtt{a} \in \Sigma$
- Wenn α und β reguläre Ausdrücke sind,
  dann sind auch $(\alpha\beta)$, $(\alpha \mid \beta)$ und $(\alpha)^*$ reguläre Ausdrücke

Die **Sprache eines regulären Ausdrucks** α wird mit $\mathbf{L}(\alpha)$ bezeichnet und rekursiv definiert:

- $\mathbf{L}(\emptyset) = \emptyset$
- $\mathbf{L}(\varepsilon) = \{\varepsilon\}$
- $\mathbf{L}(\mathtt{a}) = \{\mathtt{a}\} \quad \forall \mathtt{a} \in \Sigma$
- $\mathbf{L}((\alpha\beta)) = \mathbf{L}(\alpha) \circ \mathbf{L}(\beta)$
- $\mathbf{L}((\alpha \mid \beta)) = \mathbf{L}(\alpha) \cup \mathbf{L}(\beta)$
- $\mathbf{L}((\alpha)^*) = \mathbf{L}(\alpha)^*$

AUFGABE 1

Gegeben sind das Alphabet $\Sigma = \{a, b, c\}$ und die Sprache

$$\begin{aligned} L = \{w \in \Sigma^* \mid\ & \text{es gibt } u, v \in \Sigma^* \text{ mit } w = ubabcv \text{ und} \\ & \text{es gibt } u, v \in \Sigma^* \text{ mit } w = ucccv \text{ und} \\ & \text{es gibt kein } u \in \Sigma^* \text{ mit } w = au\}. \end{aligned}$$

Geben Sie für L einen regulären Ausdruck r mit $L = L(r)$ an.

AUFGABE 2

Beweisen Sie die folgenden Gleichungen für reguläre Ausdrücke r, s und t.

Erinnerung: $r \equiv s$ bedeutet $L(r) = L(s)$

a) $r \mid s \equiv s \mid r$

b) $(r \mid s) \mid t \equiv r \mid (s \mid t)$

c) $(rs)t \equiv r(st)$

d) $r(s \mid t) \equiv rs \mid rt$

e) $\emptyset^* \equiv \varepsilon$

f) $(r^*)^* \equiv r^*$

g) $r^* \equiv rr^* \mid \varepsilon$

h) $(\varepsilon \mid r)^* \equiv r^*$

DIE ERSETZUNGSMETHODE

Gegeben: NFA $\mathcal{M} = \langle Q, \Sigma, \delta, Q_0, F \rangle$

Gesucht: regulärer Ausdruck α mit $\mathbf{L}(\alpha) = \mathbf{L}(\mathcal{M})$

Idee:

Für jeden Zustand $q \in Q$, berechne einen regulären Ausdruck α_q für die Sprache $\mathbf{L}(\alpha_q) = \mathbf{L}(\mathcal{M}_q)$ mit $\mathcal{M}_q = \langle Q, \Sigma, \delta, \{q\}, F \rangle$

Für $Q_0 = \{q_1, q_2, \ldots, q_n\}$ gilt dann

$$\mathbf{L}(\mathcal{M}) = \bigcup_{q \in Q_0} \mathbf{L}(\alpha_q) = \mathbf{L}(\alpha_{q_1} \mid \alpha_{q_2} \mid \cdots \mid \alpha_{q_n})$$

(1) **Vereinfache den Automaten** (entferne offensichtlich unnötige Zustände)

(2) **Bestimme das Gleichungssystem**

Intuition: Beschreibe α_q in Abhängigkeit von Folgezuständen

- ▷ Für jeden Zustand $q \in Q \setminus F$: $\alpha_q \equiv \sum_{\mathtt{a} \in \Sigma} \sum_{p \in \delta(q,\mathtt{a})} \mathtt{a}\alpha_p$
- ▷ Für jeden Zustand $q \in F$: $\alpha_q \equiv \varepsilon \mid \sum_{\mathtt{a} \in \Sigma} \sum_{p \in \delta(q,\mathtt{a})} \mathtt{a}\alpha_p$

(3) **Löse das Gleichungssystem** durch Einsetzen und

Regel von Arden: Aus $\alpha \equiv \beta\alpha \mid \gamma$ mit $\varepsilon \notin \mathbf{L}(\beta)$ folgt $\alpha \equiv \beta^*\gamma$.

(4) **Gib den Ausdruck für die Sprache des NFA an**

Für $Q_0 = \{q_1, q_2, \ldots, q_n\}$ gilt dann

$$\mathbf{L}(\mathcal{M}) = \bigcup_{q \in Q_0} \mathbf{L}(\alpha_q) = \mathbf{L}(\alpha_{q_1} \mid \alpha_{q_2} \mid \cdots \mid \alpha_{q_n})$$

DYNAMISCHE ERMITTLUNG

Gegeben: NFA $\mathcal{M} = \langle Q, \Sigma, \delta, Q_0, F \rangle$

Gesucht: regulärer Ausdruck α mit $\mathbf{L}(\alpha) = \mathbf{L}(\mathcal{M})$

Ansatz:

Für jedes Paar von Zuständen $q, p \in Q$, berechne einen regulären Ausdruck $\alpha_{q,p}$ für die Sprache $\mathbf{L}(\alpha_{q,p}) = \mathbf{L}(\mathcal{M}_{q,p})$ mit $\mathcal{M}_{q,p} = \langle Q, \Sigma, \delta, \{q\}, \{p\} \rangle$

Dann gilt:

$$\mathbf{L}(\mathcal{M}) = \bigcup_{q \in Q_0} \bigcup_{p \in F} \mathbf{L}(\alpha_{q,p}) = \mathbf{L}\left(\sum_{q \in Q_0} \sum_{p \in F} \alpha_{q,p} \right)$$

- $\mathbf{L}^k[i,j] \ldots$ Sprache mit Start in q_i, Ende in q_j und nutzt nur Zwischenzustände $q_1, \ldots, q_k$
- $\alpha^k[i,j]$ zugehöriger regulärer Ausdruck

Idee: lasse immer mehr Zwischenzustände von q nach p zu

- $k = n$: nutze alle Zustände — Ergebnis ablesbar
- $k = 0$: nutze keine Zwischenzustände — $\alpha^0[i,j]$ direkt ablesbar:

Sei $\{\mathtt{a}_1, \ldots, \mathtt{a}_m\} = \{\mathtt{a} \in \Sigma \mid q_i \xrightarrow{\mathtt{a}} q_j\}$ die Menge der Beschriftungen von direkten Übergängen von q_i zu q_j.

- ▷ Falls $i \neq j$, dann ist $\alpha^0[i,j] = \mathtt{a}_1 \mid \ldots \mid \mathtt{a}_m$
- ▷ Falls $i = j$, dann ist $\alpha^0[i,j] = \mathtt{a}_1 \mid \ldots \mid \mathtt{a}_m \mid \varepsilon$

Update-Formel:

$$\alpha^{k+1}[i,j] = \alpha^k[i,j] \mid \left(\alpha^k[i,k+1](\alpha^k[k+1,k+1])^*\alpha^k[k+1,j]\right)$$

vgl. VL „Algorithmen & Datenstrukturen", Prozess-Problem im Aho-Hopcroft-Ullman-Algorithmus

AUFGABE 3

Geben Sie zu jedem der regulären Ausdrücke r_i einen NFA $\mathcal{M}_i$ mit $L(\mathcal{M}_i) = L(r_i)$ an.

a) $r_1 = (ab)^*$

b) $r_2 = a(b \mid c)a^* \mid a^*$

Wenden Sie dabei jeweils den *kompositionellen Ansatz* sowie den *expliziten Ansatz* zur Konstruktion von NFAs aus der Vorlesung an.

AUFGABE 4

Entwickeln Sie für die Sprache L über dem Alphabet $\Sigma = \{a, b, c\}$ einen regulären Ausdruck r mit $L = L(r)$. Für alle Wörter $w \in L$ gilt:

- w enthält aaa.
- w endet mit c.
- Die Anzahl der b in w ist gerade.

ÄQUIVALENZ VON ZUSTÄNDEN & QUOTIENTENAUTOMAT

DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle \rightsquigarrow$ DFA $\mathcal{M}_q = \langle Q, \Sigma, \delta, q, F \rangle$

Zwei Zustände $p, q \in Q$ sind $\mathcal{M}$-**äquivalent** ($p \sim_{\mathcal{M}} q$), falls

$$\mathbf{L}(\mathcal{M}_p) = \mathbf{L}(\mathcal{M}_q)$$

Die Äquivalenzklasse eines Zustands $q \in Q$ ist

$$[q]_\sim = \{p \in Q \mid q \sim p\}$$

Für eine Menge $P \subseteq Q$ schreiben wir $P/_\sim$ für den Quotienten von P und $\sim$:

$$P/_\sim = \{[p]_\sim \mid p \in P\}$$

- $\sim$ ist eine Äquivalenzrelation (reflexiv, symmetrisch, transitiv)
- Äquivalenzklassen sind disjunkt und partitionieren Q

Quotientenautomat

Idee: Verschmelzen von äquivalenten Zuständen

Für einen DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ mit totaler Übergangsfunktion ist der Quotientenautomat $\mathcal{M}/_\sim$ gegeben durch

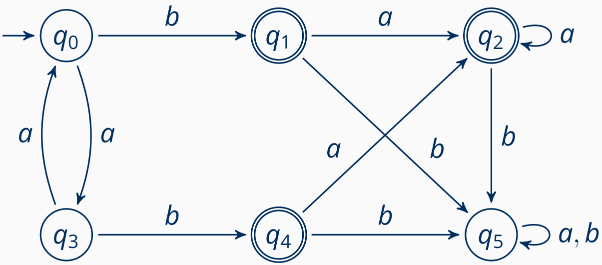
$$\mathcal{M}/_\sim = \langle Q/_\sim, \Sigma, \delta_\sim, [q_0]_{\sim_\mathcal{M}}, F/_\sim \rangle$$

wobei gilt:

- $Q/_\sim = \{[q]_\sim \mid q \in Q\}$
- $\delta_\sim([q]_\sim, \mathtt{a}) = [\delta(q, \mathtt{a})]_\sim$
- $F/_\sim = \{[q]_\sim \mid q \in F\}$

AUFGABE 5

Berechnen Sie für folgenden DFA
$\mathcal{M} = (\{q_0, q_1, q_2, q_3, q_4, q_5\}, \{a, b\}, \delta, q_0, \{q_1, q_2, q_4\})$ mit δ:



die Äquivalenzrelation $\sim_{\mathcal{M}}$, und geben Sie den Quotientenautomaten $\mathcal{M}/_{\sim}$ an.