SS 2015

Zentralübung zur Vorlesung Theoretische Informatik

Dr. Werner Meixner

Fakultät für Informatik
TU München

http://www14.in.tum.de/lehre/2015SS/theo/uebung/

30. April 2015

ZÜ III

Übersicht:

1. Übungsbetrieb Fragen, Probleme?
2. Thema Reguläre Ausdrücke
3. Vorbereitung TA Blatt 3

1. Fragen, Anregungen?

Aktuelle Fragen?

2. Thema: Reguläre Ausdrücke

Reguläre Ausdrücke stellen Elemente einer Mengenalgebra dar.

Die Operationen der Mengenalgebra sind die Konkatenation, Vereinigung („$|$"), Sternbildung („*") und die Operationen $\emptyset$, ϵ, sowie eine endliche Anzahl von Konstanten,
d.h. nullstelligen Operationen $a, b, \ldots$, die eineindeutig den Elementen eines Alphabets Σ entsprechen.

Reguläre Ausdrücke sind also nichts anderes als algebraische Ausdrücke ohne Variable über einer Mengenalgebra.

Reguläre Ausdrücke gehorchen den Gesetzen algebraischer Ausdrücke.

Die Sprache der regulären Ausdrücke werden wir nicht vollständig formalisieren, insbesondere greifen wir auf bekannte Gesetze der Klammerbildung bei algebraischen Ausdrücken zurück.

3. Vorbereitung TA Blatt 3

3.1 VA 1

Studieren Sie die Definition des Begriffs des regulären Ausdrucks und beantworten Sie kurz die folgenden Fragen:

1. Welche Mengen stellen die Ausdrücke $\emptyset$ bzw. ϵ bzw. a, b dar?
2. Geben sie einen regulären Ausdruck an, der eine Sprache A über dem Alphabet $\Sigma = \{0, 1\}$ mit den Eigenschaften $01 \in A$ und $A^*A = A$ darstellt.
3. Finden Sie einen regulären Ausdruck über dem Alphabet $\Sigma = \{0, 1\}$, der die Menge aller Wörter beschreibt, die mit 00 beginnen und in denen 1 genau dreimal vorkommt.

1. Welche Mengen stellen die Ausdrücke $\emptyset$ bzw. ϵ bzw. a, b dar?

Lösung

Wir benützen eine drucktechnische Unterscheidung durch Fettdruck.

$L(\boldsymbol{\emptyset}) = \emptyset$,

$L(\boldsymbol{\epsilon}) = \{\epsilon\}$,

$L(\boldsymbol{a}) = \{a\}$, $L(\boldsymbol{b}) = \{b\}$.

2. Geben sie einen regulären Ausdruck an, der eine Sprache A über dem Alphabet $\Sigma = \{0, 1\}$ mit den Eigenschaften $01 \in A$ und $A^*A = A$ darstellt.

Lösung

$A = \{0, 1\}^*$ erfüllt die Anforderungen an die Sprache. (Beweis?)

Sei $\alpha = (\mathbf{0}|\mathbf{1})^*$ ein regulärer Ausdruck.
Es gilt $A = L(\alpha)$. (Beweis?)

Die mengentheoretisch kleinste Sprache, für die die Anforderungen erfüllt sind, ist

$A' := \{01\}^*\{01\}$, d.h. $A' = L((\mathbf{01})^*(\mathbf{01}))$.

Beweis

Jede Sprache A, für die die Anforderungen erfüllt sind, enthält A':

Falls $A^*A = A$ und $01 \in A$, folgt $\{01\} \subseteq A$ mithin

$\{01\}^*\{01\} \ \subseteq \ A$

d.h. $A' \subseteq A$.

Umgekehrt gelten für $A' = \{01\}^*\{01\}$ die Anforderungen,
d.h. die Implikationen

$01 \in A'$ und $A'^*A' \subseteq A'$.

Algebraischer Beweis von $A'^*A' = A'$:

$$\begin{array}{rcll} A'^*A' & = & (\{01\}^*\{01\})^*\{01\}^*\{01\} & \\ & = & (\{01\}^*)^*\{01\}^*\{01\}^*\{01\} & \text{wg. } (XY)^* = X^*Y^* \\ & = & \{01\}^*\{01\}^*\{01\}^*\{01\} & \text{wg. } (X^*)^* = X^* \\ & = & \{01\}^*\{01\} & \text{wg. } (XY)^* = X^*Y^* \\ & = & A'\,. & \end{array}$$

3. Finden Sie einen regulären Ausdruck über dem Alphabet $\Sigma = \{0, 1\}$, der die Menge aller Wörter beschreibt, die mit 00 beginnen und in denen 1 genau dreimal vorkommt.

Lösung

Eine Lösung ist $000^*10^*10^*10^*$.

3.2 VA 2

Reguläre Ausdrücke α, β heißen äquivalent, i.Z. $\alpha \equiv \beta$, genau dann, wenn $L(\alpha) = L(\beta)$ gilt.

Beweisen Sie für alle regulären Ausdrücke α und β:
$(\alpha\beta)^*\alpha \equiv \alpha(\beta\alpha)^*$.

Die Äquivalenz $r \equiv s$ von regulären Ausdrücken r und s ist als Gleichheit der zugeordneten Sprachen $L(r)$ und $L(s)$ definiert. Für diese Äquivalenzgleichungen gelten einfache Regeln.

Wir zeigen zunächst für zwei beliebige Sprachen A und B die Gleichung

$$(AB)^n A = A (BA)^n \tag{1}$$

per Induktion.

Induktion:

- $n = 0$: Es gilt $(AB)^0\, A = A = A\,(BA)^0$.
- $n \to n + 1$: Es gilt

$$\begin{aligned}(AB)^{n+1}\, A &= AB\,(AB)^n\, A \\ &= ABA\,(BA)^n \qquad\qquad \text{(mit Ind.hyp.)} \\ &= A\,(BA)^{n+1}\ .\end{aligned}$$

Nun wird die Gleichung definitionsgemäß semantisch mit den zugeordneten Sprachen interpretiert und die für Mengen von Wörtern, d.h. Sprachen, geltenden Rechenregeln angewandt:

$$
\begin{aligned}
L\left((\alpha\beta)^*\alpha\right) &= \left(L(\alpha)L(\beta)\right)^* L(\alpha) \\
&= \bigcup_{n\geq 0} \left(L(\alpha)L(\beta)\right)^n L(\alpha) \\
&= \bigcup_{n\geq 0} L(\alpha)\left(L(\beta)L(\alpha)\right)^n \qquad \text{(mit Gl. 1)} \\
&= L(\alpha)\left(L(\beta)L(\alpha)\right)^* \\
&= L\left(\alpha(\beta\alpha)^*\right) .
\end{aligned}
$$

3.3 VA 3

Wann genau ist die von einem endlichen Automaten erzeugte Sprache endlich?

Beantworten Sie diese Frage anhand der Länge der Pfade, die es in dem Übergangsgraphen eines Automaten gibt.

Antwort

Genau dann, wenn es keinen erreichbaren und „produktiven“ (d.h., von dort aus ist ein Endzustand erreichbar) Zustand gibt, der auf einem Kreis liegt.

3.4 VA 4

Wahr oder falsch? Begründen Sie Ihre Antwort!

1. Zu jedem ϵ-NFA $N = (Q, \Sigma, \delta, \{q_0\}, F)$ gibt es einen äquivalenten (ϵ-freien) NFA $N' = (Q', \Sigma, \delta', \{q_0\}', F')$, so dass $|Q'| \leq |Q|$ gilt.
2. Für jeden NFA $N = (Q, \Sigma, \delta, \{q_0\}, F)$ gilt: Wenn $F = Q$, dann ist $L(N) = \Sigma^*$.
3. $a(ab)^*(ba)^*b \equiv a(ab|ba)^*b$.
4. Wenn $L \subseteq \Sigma^*$ regulär ist und $\Gamma \subseteq \Sigma$, dann ist $L' = \{w \in L \mid w \text{ enthält nur Zeichen aus } \Gamma\}$ ebenfalls regulär.

1. Zu jedem ϵ-NFA $N = (Q, \Sigma, \delta, \{q_0\}, F)$ gibt es einen äquivalenten (ϵ-freien) NFA $N' = (Q', \Sigma, \delta', \{q_0\}', F')$, so dass $|Q'| \leq |Q|$ gilt.

Lösung

Wahr! Die Konstruktion in der Vorlesung liefert $Q = Q'$.

2. Für jeden NFA $N = (Q, \Sigma, \delta, \{q_0\}, F)$ gilt:
   Wenn $F = Q$, dann ist $L(N) = \Sigma^*$.

Lösung

Falsch!
Gegenbsp.: $\Sigma = \{a\}, Q = F = \{q_0\}, \delta(q_0, a) = \emptyset$.
Dann ist $L(N) = \{\epsilon\}$.

3. $a(ab)^*(ba)^*b \equiv a(ab|ba)^*b$.

Lösung

Falsch!
$w = aabbaabb \in L(a(ab|ba)^*b)$, aber $w \notin L(a(ab)^*(ba)^*b)$.

4. Wenn $L \subseteq \Sigma^*$ regulär ist und $\Gamma \subseteq \Sigma$,
dann ist $L' = \{w \in L \mid w$ enthält nur Zeichen aus $\Gamma\}$
ebenfalls regulär.

Lösung

Wahr! Abschluss unter Schnitt: $L' = L \cap \Gamma^*$.