Grundlagen der theoretischen Informatik SoSe 2018

Anton Oehler, Jona Rex

Abgabe am 4. Mai 2018 bis 10 Uhr in die Tutorenfächer

Aufgabe 1 Reguläre Ausdrücke I *10 Punkte*

(a) Geben Sie reguläre Ausdrücke für die folgenden Sprachen an. Erklären Sie Ihren Ausdruck in einem Satz, und geben Sie das zu Grunde liegende Alphabet an. Erklären Sie ggf. auch alle Annahmen, die Sie gemacht haben.

(i) Summen von positiven dezimalen Festkommazahlen. In der Sprache sollen also z.B. die Zeichenketten 3.14 oder auch $3 + 4.2 + 7 + 1$ vorkommen.

$$\begin{aligned} \Sigma &= \{0, \dots, 9, \text{`.`}, +\} \\ \text{Sei } d &= \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \\ \text{und } f &= d \cup dd \\ R &= d^+\{\text{`.`} \circ f, \epsilon\}\ (+ \circ d^+\{\text{`.`} \circ f, \epsilon\})^* \end{aligned}$$

Eine Summe von positiven dez. Festkommazahlen besteht aus einer ersten Zahl und optional einer beliebig langen Folge von '+' und weiteren Zahlen. Jede Zahl setzt sich aus mindestens einer Ziffer zusammen, optional gefolgt von einem Punkt (hier der Sichtbarkeit halber '.') und einer oder zwei weiteren Ziffern.

Annahme: Festkommazahlen beschränken sich auf 2 Nachkommastellen, sonst ist $f = d^+$ und auch der Teil vor dem Komma darf mit 0 beginnen.

(ii) KFZ-Nummernschilder, bestehend aus ein bis drei Großbuchstaben gefolgt von einem Bindestrich, dann ein bis zwei Großbuchstaben und zum Schluss ein bis vier Ziffern. Die Länge der Nummernschilder insgesamt beträgt nicht mehr als acht Zeichen.

$$\begin{aligned} \Sigma &= \{A, \dots, Z, 0, \dots, 9, -\} \\ \text{Sei } a &= \{A, \dots, Z\} \\ \text{und } d &= \{0, \dots, 9\} \\ R &= (a - (a \cup aa)(d \cup dd \cup ddd \cup dddd)) \cup \\ &\quad (aa - a(a \cup d \cup \epsilon)(d \cup dd \cup ddd)) \cup \\ &\quad (aaa - a(a \cup d \cup \epsilon)(d \cup dd) \end{aligned}$$

Wir betrachten 3 Fälle: (hier: Buchstabe = Großbuchstabe)

- 1 Buchstabe am Anfang, gefolgt von einem Bindestrich und 1-2 Buchstaben und 1-4 Zahlen
- 2 Buchstaben am Anfang, gefolgt von einem Bindestrich und einem Buchstaben. Dann folgt [entweder ein Buchstabe, eine Zahl oder nichts] und danach 1-3 Zahlen
- 3 Buchstaben am Anfang, ein Bindestrich, gefolgt von einem Buchstaben, gefolgt von [entweder einem Buchstaben, einer Zahl oder nichts] und anschließend 1-2 Zahlen

Annahme: Der Bindestrich zählt auch als Zeichen, nicht nur Buchstaben und Zahlen.

(iii) Die Menge aller Binärzahlen, die durch vier teilbar sind. Dabei sind keine überflüssigen führenden Nullen erlaubt.

$$R = 0 \cup (1(0 \cup 1)^* \cup \epsilon)100$$

Entweder 0, 100 ($\epsilon \circ 100$) oder eine Zahl die mit 1 beginnt, gefolgt von beliebigen Kombinationen von 0 und 1, gefolgt von 100.

(b) Beschreiben Sie in deutscher Sprache die durch die folgenden regulären Ausdrücke charakterisierten Mengen. Ihre Beschreibungen sollen die Form haben:

Die Menge aller Wörter über dem Alphabet $\{0, 1\}$ …. "

Dabei soll … durch maximal acht Wörter ersetzt werden.

(i) $0^*(0^*10^*10^*)^*$ Beliebige Folge mit gerader Anzahl an Einsen

(ii) $(00 \cup 11 \cup (01 \cup 10)(00 \cup 11)^*(01 \cup 10))^*$.

Aufgabe 2 Reguläre Ausdrücke II *10 Punkte*

Zwei reguläre Ausdrücke α, β heißen *äquivalent*, geschrieben $\alpha \sim \beta$, genau dann, wenn sie die gleiche Sprache repräsentieren, also $L(\alpha) = L(\beta)$ ist. Zeigen Sie:

(a) $(\alpha \cup \beta)\gamma \sim \alpha\gamma \cup \beta\gamma$
Zu zeigen:

$$\begin{aligned} L((\alpha \cup \beta)\gamma) &= L(\alpha\gamma \cup \beta\gamma) \\ L(\alpha \cup \beta) \circ L(\gamma) &= L(\alpha\gamma) \cup L(\beta\gamma) \\ \{\alpha, \beta\} \circ \{\gamma\} &= \{\alpha \circ \gamma\} \cup \{\beta \circ \gamma\} \\ \{\alpha \circ \gamma, \beta \circ \gamma\} &= \{\alpha \circ \gamma, \beta \circ \gamma\} \qquad \square \end{aligned}$$

(b) $(\alpha^*)^* \sim \alpha^*$
Zu zeigen:

$$\begin{aligned} L((\alpha^*)^*) &= L(\alpha^*) \\ L(\alpha^*) &= \{\epsilon, \alpha, \alpha\alpha, \alpha\alpha\alpha, \dots\} \end{aligned}$$

Konkateniert man nun für $L((\alpha^*)^*)$ alle Wörter aus $\{\epsilon, \alpha, \alpha\alpha, \alpha\alpha\alpha, \dots\}$ so kann dadurch nur wieder $\{\epsilon, \alpha, \alpha\alpha, \alpha\alpha\alpha, \dots\}$ entstehen:

$$\begin{aligned} \epsilon \circ \epsilon &= \epsilon \\ \epsilon \circ \alpha &= \alpha \\ \epsilon \circ \alpha\alpha &= \alpha\alpha \\ &\dots \\ \alpha \circ \epsilon &= \alpha \\ \alpha \circ \alpha &= \alpha\alpha \\ &\dots \end{aligned}$$

Also ist $L(\alpha^*)^2 = L(\alpha^*)$ und nach gleichem Schema auch

$$\forall k \in \mathbb{N}, k > 0 : L(\alpha^*)^k = L(\alpha^*)$$

und für $k = 0 : L(\alpha^*)^k = L(\alpha^*)^0 = \{\epsilon\}$. Daraus folgt:

$$\begin{aligned}
L(\alpha^*)^* &= \bigcup_{k=0}^{\infty} L(\alpha^*)^k \\
&= L(\alpha^*)^0 \, \cup \bigcup_{k=1}^{\infty} L(\alpha^*)^k \\
&= \{\epsilon\} \, \cup \bigcup_{k=1}^{\infty} L(\alpha^*)^k \\
&= \{\epsilon\} \, \cup \bigcup_{k=1}^{\infty} L(\alpha^*) \\
&= \{\epsilon\} \, \cup L(\alpha^*) \\
&= L(\alpha^*) \qquad \square
\end{aligned}$$

(c) $(\alpha \cup \beta)^* \sim (\alpha^* \beta^*)^*$;

(d) $\alpha\emptyset \sim \emptyset$
Zu zeigen:

$$\begin{aligned}
L(\alpha\emptyset) &= L(\emptyset) \\
\{w \circ v | w \in \{\alpha\} \wedge v \in \emptyset\} &= \emptyset \\
\emptyset &= \emptyset \qquad \square
\end{aligned}$$

(e) $\alpha\emptyset^* \sim \alpha$
Zu zeigen:

$$\begin{aligned}
L(\alpha\emptyset^*) &= L(\alpha) \\
L(\alpha) \circ L(\emptyset^*) &= L(\alpha)
\end{aligned}$$

Per Definition ist $\emptyset \circ \emptyset = \emptyset$ und $\emptyset \cup \emptyset = \emptyset$. Somit ist

$$\begin{aligned}
\emptyset^* &= \emptyset^0 \cup \emptyset^1 \cup \emptyset^2 \ldots \\
&= \emptyset^0 \cup \emptyset \cup \emptyset \ldots \\
&= \{\epsilon\}
\end{aligned}$$

Also ist

$$L(\alpha) \circ L(\emptyset\} = L(\alpha) \circ \{\epsilon\} = L(\alpha) \qquad \square$$

Aufgabe 3 Reguläre Sprachen *10 Punkte*

Eine Sprache heißt *regulär*, falls sie sich durch einen regulären Ausdruck darstellen lässt.

(a) Zeigen Sie: Jede endliche Sprache ist regulär.
Antwort:
Eine endliche Sprache besteht nur aus endlich vielen Wörtern. Eine (primitive) Darstellung dieser Sprache ist, alle möglichen Wörter miteinander zu vereinigen.

(b) Sei Σ ein Alphabet und $w = \sigma_1\sigma_2 \ldots \sigma_n$ ein Wort über Σ. Die *Umkehrung* von w, w^R, ist definiert als $w^R = \sigma_n\sigma_{n-1} \ldots \sigma_1$. So ist zum Beispiel $\text{haus}^R = \text{suah}$, $\text{blatt}^R = \text{ttalb}$, $\text{a}^R = \text{a}$ und $\varepsilon^R = \varepsilon$. Die Umkehrung einer Sprache $L \subseteq \Sigma^*$ ist definiert als $L^R = \{w^R \mid w \in L\}$.
Zeigen Sie: Ist L eine reguläre Sprache, so ist auch L^R regulär. Geben Sie dazu einen detaillierten Beweis, der strukturelle Induktion über den regulären Ausdruck für L verwendet.
Antwort:
Sei w ist Wort über Σ mit $w = o_1o_2 \ldots o_n$ und
w^R seine Umkehrung mit $o_no_{n-1} \ldots o_1$.
Sei die Umkehrung einer Sprache $L \subseteq \Sigma^*$ definiert als $L^R = \{w^R | w \in L\}$.

Voraussetzung: L ist regulär $\Rightarrow$ L^R ist regulär

Anfang: mit $L = \emptyset$:
Aus der Definition von $L^R = \{w^R | w \in L\}$ folgt:

$$L = \emptyset \Leftrightarrow L^R = \emptyset$$

mit $L = \{\epsilon\}$:
Per Definition ist $\epsilon^R = \epsilon$. Also gilt:

$$L = \{\epsilon\} \Leftrightarrow L^R = \{\epsilon\}$$

Schritt: Laut *IV.* gilt:
Wenn L regulär, dann ist auch L^R regulär.
Sei nun $w = o_1o_2 \ldots o_n$ ein beliebiges Wort der Länge n in L und
$w^R = o_no_{n-1} \ldots o_1$ seine Umkehrung in L^R.
Zu zeigen: für $a \in \Sigma$ gilt:

- $(w \circ a) \in L'$,(L' reg.) $\Rightarrow (w \circ a)^R \in L'^R$,($L'^R$ reg.)

$$\begin{aligned} &(w \circ a) = o_1o_2 \ldots o_na \\ \Rightarrow &(w \circ a)^R = ao_no_{n-1} \ldots o_2o_1 \\ \Rightarrow &L'^R \text{ ist regulär} \end{aligned}$$

- $(w \cup a) \in L'$,(L' reg.) $\Rightarrow (w \cup a)^R \in L'^R$,($L'^R$ reg.)

$$\begin{aligned} &(w \cup a) = \{w, a\} \\ \Rightarrow &(w \cup a)^R = \{w^R, a^R\} \\ \Rightarrow &(w \cup a)^R = \{w^R, a\} \end{aligned}$$

Dass $w^R \in L'^R$ lässt sich wie zuvor zeigen, und da $a^R = a$ (Definition), ist $(w \cup a)^R \in L'^R$ und L'^R regulär.