Abschlusseigenschaften regulärer Sprachen

Die Klasse der regulären Sprachen ist abgeschlossen unter

a) ✓ Mengenoperationen $L \cup M$, $L \cap M$, $L \setminus M$, $\underbrace{\Sigma^* \setminus L}_{\text{Komplement}}$ ✓

b) ✓ Multiplikation (Verkettung) LM ✓, $*$-Operation ✓ L^*

c) ✓ Spiegelung (Umkehrung) L^R

d) Homomorphismus $h(L)$

e) inversem Homomorphismus $h^{-1}(L)$.

D.h.: Wenn L und M reguläre Sprachen sind, dann ist auch $L \cup M$, $L \cap M$, ... eine reguläre Sprache.

<u>Satz</u>: Die folgenden Möglichkeiten beschreiben die gleiche Klasse von Sprachen (die regulären Sprachen):

1. reguläre Ausdrücke (r.A.)
2. deterministische endliche Automaten (DEA)
3. nichtdeterministische endliche Automaten (NEA) (mit ε-Übergängen)

a) $L \cup M$ b) LM, L^* : r.A.

a) $\overline{L} := \Sigma^* \setminus L$ → Übung 25

$L \cap M = \overline{\overline{L} \cup \overline{M}}$ (de Morgan'sches Gesetz)

a) $L \setminus M = L \cap \overline{M}$

$L \cup M$ → Übung 17
$L \cap M$
$L \setminus M$
} Produktautomat

c) Abschluss unter Spiegelung. (r. A.)

Wenn es einen regulären Ausdruck für eine Sprache L gibt, dann gibt es auch einen regulären Ausdruck für L^R.

$$(L \cup M)^R = L^R \cup M^R$$
$$(LM)^R = M^R . L^R$$
$$(L^*)^R = (L^R)^* \qquad (L^i)^R = (L^R)^i$$
$$(\{a\})^R = \{a\}, \quad a \in \Sigma$$
$$\emptyset^R = \emptyset$$
$$\{\varepsilon\}^R = \{\varepsilon\}$$

Formale Definition eines regulären Ausdrucks über Σ.

- $\emptyset$ ist r.A.
- ε ist r.A.
- a ist r.A. ($a \in \Sigma$)
- Wenn R und S r.A. sind, dann ist $(R \cup S)$ r.A.
- .. dann ist $(R \cdot S)$ r.A.
- Wenn R ein r.A. ist, dann ist (R^*) r.A

nach strengen und formalistischen (aber einfachen) Regeln

r. A. ist nur, was gemäß diesen Regeln aufgebaut werden kann.

Bsp.

$$\underbrace{(((((0 \cdot 1) \cdot 1)^*) \cup \varepsilon) \cup (0 \cdot \emptyset))}$$

konventionelle Schreibweise:

$$\{011\}^* \cup \{\varepsilon\} \cup \{0\} \cdot \emptyset$$

Diese Definition definiert einen regulären Ausdruck als eine <u>Zeichenkette</u> über dem Alphabet $\{ (,), ^*, \cdot, \cup, \varepsilon, \emptyset \} \cup \Sigma$

Induktive Definition der von einem regulären Ausdruck R dargestellten Sprache $L(R)$:

$L(\emptyset) = \emptyset$

$L(\varepsilon) = \{\varepsilon\}$

$L(a) = \{a\} \quad a \in \Sigma$

$L(\underbrace{(R \cup S)}) = L(R) \cup L(S)$

$L(\underbrace{(R \cdot S)}) = L(R) \cdot L(S)$

$L(\underbrace{(R^*)}) = (L(R))^*$

Wenn es einen regulären Ausdruck S für eine Sprache L gibt, dann gibt es auch einen regulären Ausdruck für L^R.

Beweis mit struktureller Induktion nach S.

$L(S) = L$

zum Vergleich:
vollständige Induktion nach n

1) $n = 0$ 2) $n \to n+1$

1) $S = \emptyset$: $L(\underline{\emptyset}) = L^R$

2) $S = \varepsilon$: $L(\underline{\varepsilon}) = L^R$

3) $S = a \in \Sigma$: $L(\underline{a}) = L^R$

4) S ist von der Form $(U \cup V)$ $L(S) = L = L(U) \cup L(V)$

(*) $L(\underline{?}) = L^R = (L(U))^R \cup (L(V))^R$

nach Induktionsvoraussetzung gibt es

- einen Ausdruck S_1 mit $L(S_1) = (L(U))^R$
- einen Ausdruck S_2 mit $L(S_2) = (L(V))^R$

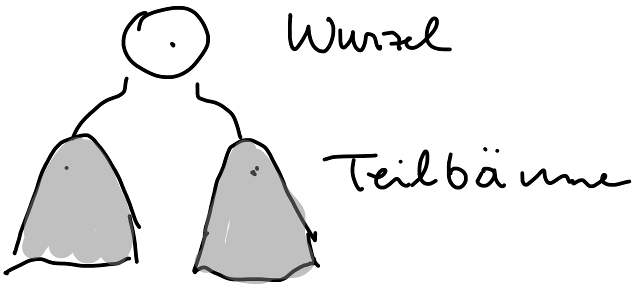
Der gesuchte Ausdruck (*) ist $(S_1 \cup S_2)$

5) S ist von der Form $(U \cdot V)$ $L(S) = L = L(U)\, L(V)$

(*) $L(?) = L^R = \underbrace{(L(V))^R}_{L(S_2)} \cdot \underbrace{(L(U))^R}_{L(S_1)}$

Der gesuchte Ausdruck (*) ist $(S_2 \cdot S_1)$

6) S ist von der Form (U^*) .... $(S_1{}^*)$



$$\left(\{01,011\}\cdot\{110\}^*\cdot\left(\left(\{\varepsilon\}\cup\{10\}^*\right)\{1100\}\right)^*\right)^R$$

$$=\left(\left(\left(\{\varepsilon\}\cup\{10\}^*\right)\{1100\}\right)^*\right)^R\cdot\left(\{110\}^*\right)^R\cdot\left(\{01,011\}\right)^R$$

$$\left(\left(\left(\{\varepsilon\}\cup\{10\}^*\right)\cdot\{1100\}\right)^R\right)^*\ \underbrace{\left(\{110\}^R\right)^*}_{\{011\}^*}\ \{10,110\}$$

$$\left(\{1100\}^R\cdot\left(\{\varepsilon\}\cup\{10\}^*\right)^R\right)^*$$

$$\left(\{0011\}\cdot\left(\{\varepsilon\}\cup\{\underset{01}{\cancel{10}}\}^*\right)\right)^*\cdot\{011\}^*\cdot\{10,110\}$$

Homomorphismen

Ein Homomorphismus $h:\Sigma^*\to(\Sigma')^*$ ist charakterisiert durch die Eigenschaft

Alphabet kann verschieden sein.

$$\forall x,y\in\Sigma^*:\quad h(x\cdot y)=h(x)\cdot h(y)$$

Produkt in $(\Sigma')^*$

„verträglich mit •"

Produkt in Σ^*

1) Setze $x = y = \varepsilon$:

$$\underbrace{h(\varepsilon \cdot \varepsilon)}_{} = h(\varepsilon) \cdot \overbrace{h(\varepsilon)}^{u}$$

$$h(\varepsilon) = h(\varepsilon) \cdot h(\varepsilon) \qquad u = u \cdot u$$

$$\boxed{h(\varepsilon) = \varepsilon}$$

2) $x = x_1 x_2 \dots x_n \in \Sigma^*$, $x_i \in \Sigma$

$$h(x) = h(x_1) h(x_2) \dots h(x_n)$$

- Ein Homomorphismus ist durch seine Werte auf den Symbolen $a \in \Sigma$ festgelegt.
- Für diese Werte können beliebige Wörter $h(a) \in (\Sigma')^*$ frei gewählt werden.

Bsp. $h(a) = 01010010$
$h(b) = 01010011$
$h(c) = 01010100$
$\vdots$

Code: $h: \{a, b, c \dots\}^* \to \{0, 1\}^*$

Bsp. $h(1) = 1$
$h(2) = \varepsilon$
$h(3) = 3$

$h(1232133) = 13133$

löscht alle Zweien

d) Homomorphismus angewendet auf eine Sprache L:

$$h(L) = \{ h(x) \mid x \in L \} \subseteq (\Sigma')^* \qquad \text{r.A. !!}$$

Wende h auf jedes Symbol $\in \Sigma$ im regulären Ausdruck R an.

Formal: $L(h(R)) = h(L(R))$

strukturelle Induktion

inverser Homomorphismus

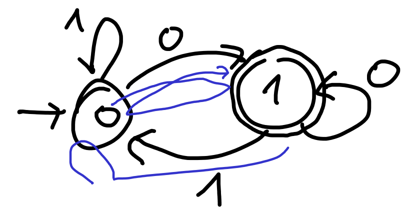
$$h^{-1}(\underset{\subseteq (\Sigma')^*}{M}) = \{ x \in \underline{\Sigma^*} \mid h(x) \in M \}$$

DEA!

$h(a) = 010$
$h(b) = 11$
$h(c) = 000$

$$h(abac) = \underbrace{010}\,\underbrace{11}\,\underbrace{010}\,\underbrace{000} \qquad q \to \delta^*(q, 010)$$

$\Sigma = \{a, b, c\}$ $\quad \Sigma' = \{0, 1\}$



$A = (Q, \Sigma', q_0, \delta, F)$

$L(A) = M$

$\delta : Q \times \Sigma' \to Q$

$x = x_1 x_2 \dots x_n \in \Sigma^*$

Für jeden Buchstaben x_i : Führe die Übergänge für $h(x_i)$ im DEA durch.

$\bar{A} = (Q, \Sigma, q_0, \bar{\delta}, F)$ $\qquad \bar{\delta} : Q \times \Sigma \to Q$

$$\bar{\delta}(q, x) := \delta^*(q, \underbrace{h(x)}_{\in (\Sigma')^*})$$

$x \in L(\bar{A}) \overset{?}{\Leftrightarrow} h(x) \in L(A)$ ✓

$x \in L(\bar{A}) \Leftrightarrow \bar{\delta}^*(q_0, x) \in F$

$\bar{\delta}(\dots \bar{\delta}(\bar{\delta}(q_0, x_1), x_2) \dots, x_n) \in F$

$\delta^*(\dots \delta^*(\delta^*(q_0, h(x_1)), h(x_2)), \dots, h(x_n)) \in F$

(*) $\downarrow$

$$\delta^*(q_0, \underbrace{h(x_1) h(x_2) \dots h(x_n)}_{h(x_1 \dots x_n) = h(x)}) \in F$$

$\Leftrightarrow \delta^*(q_0, h(x)) \in F \Leftrightarrow h(x) \in L(A)$

(*)

$$\delta^*(\underbrace{\delta^*(q, y)}_{q'}, z) = \delta^*(q, yz)$$