Wörter und Sprachen

Im Zusammenhang mit formalen Sprachen wird jede nicht-leere endliche Menge als *Alphabet* bezeichnet. Die Elemente eines Alphabets werden *Symbole* genannt.

Ein *Wort* über einem Alphabet Σ ist eine endliche Folge von Symbolen aus Σ.

Das *leere Wort* ist die aus 0 Symbolen bestehende Folge und wird mit ε bezeichnet.

$w = \varepsilon$

Die Menge aller Wörter über einem Alphabet Σ wird mit Σ^* bezeichnet. Ferner ist $\Sigma^+ = \Sigma^* - \{\varepsilon\}$.

Eine *Sprache* über einem Alphabet Σ ist eine Teilmenge von Σ^*.

Für ein Symbol σ bezeichnen wir mit $|w|_\sigma$ die Anzahl der Vorkommen von σ in w.

Seien x und y Wörter über dem gleichen Alphabet. Dann bezeichnet $x \circ y$ das Wort, das durch Hintereinanderschreiben von x und y entsteht. Diese Operation wird als *Konkatenation* oder auch als Verkettung bezeichnet. Statt $x \circ y$ schreiben wir meist einfach xy.

Für ein Wort w bezeichnet w^R das Wort w in umgekehrter Symbolfolge:

Deterministische endliche Automaten

Definition:
Ein deterministischer endlicher Automat ist ein 5-Tupel $(K, \Sigma, \delta, s, F)$, *wobei gilt:*

- K *ist eine endliche Menge von Zuständen,*
- Σ *ist ein Alphabet,*
- $s \in K$ *ist der Startzustand,*
- $\delta : K \times \Sigma \mapsto K$ *ist die Überführungsfunktion oder auch Übergangsfunktion, und*
- $F \subseteq K$ *ist die Menge der Endzustände.*

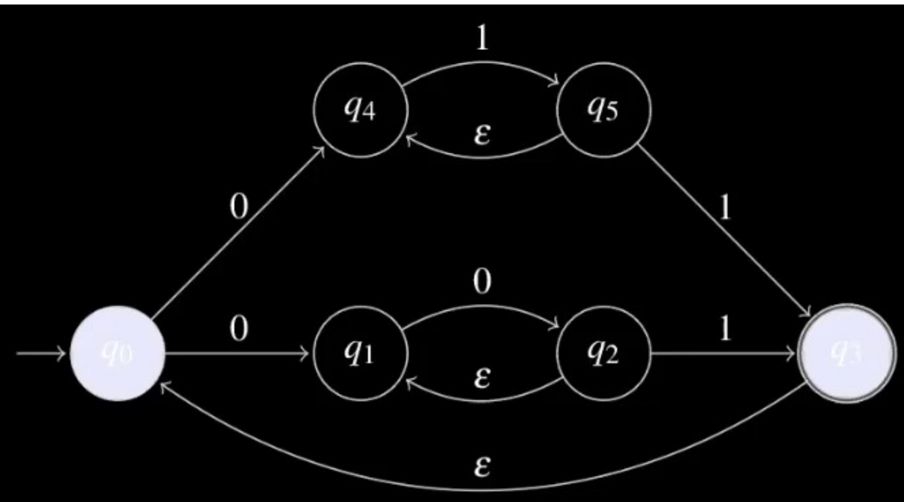
Definition:
Eine Sprache L *heißt regulär, wenn es einen endlichen Automaten* M *gibt mit* $L = \mathcal{L}(M)$.

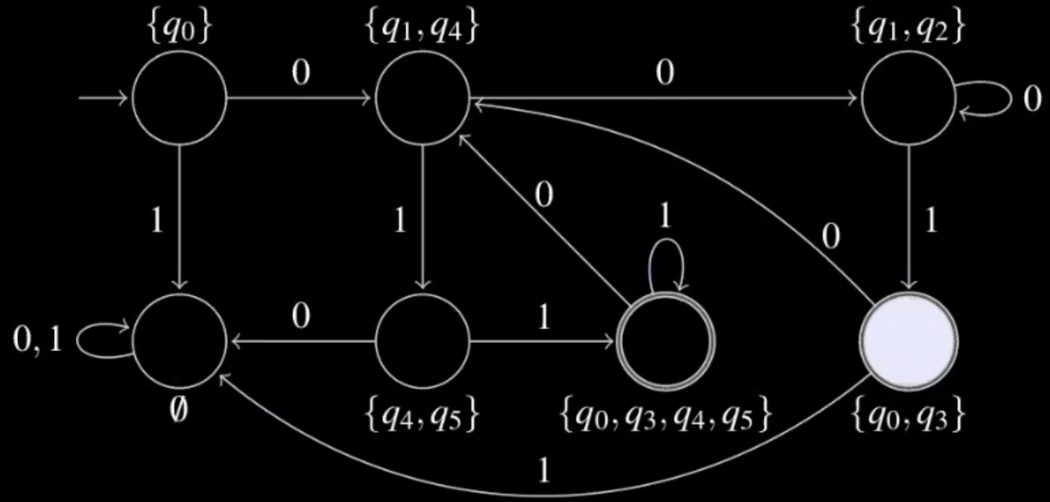
Satz:
Die Klasse der von nichtdeterministischen endlichen Automaten akzeptierten Sprachen ist abgeschlossen unter

(a) *Vereinigung,*
(b) *Konkatenation und*
(c) *Kleene-Star.*

Beweis Aequivalenz NEA und DEA

)01





Satz:

Die Klasse der von endlichen Automaten akzeptierten Sprachen ist abgeschlossen unter

(a) Vereinigung, D N
(b) Konkatenation, N
(c) Kleene Star, N
(d) Komplement und D
(e) Schnitt. D

Satz:

Die Klasse der von regulären Ausdrücken erzeugten Sprachen ist abgeschlossen unter

(a) Vereinigung,
(b) Konkatenation und
(c) Kleene Star.

Jeder reguläre Ausdruck repräsentiert eine Sprache:
Sei $\mathcal{L}(\alpha)$ die durch den Ausdruck α repräsentierte Sprache.
$\mathcal{L}$ ist wie folgt definiert:

(1) $\mathcal{L}(a) = \{a\}$ für alle $a \in \Sigma$.

(2) $\mathcal{L}(\emptyset) = \emptyset$.

(3) Falls α und β reguläre Ausdrücke sind, dann ist $\mathcal{L}((\alpha\beta)) = \mathcal{L}(\alpha)\mathcal{L}(\beta)$.

(4) Falls α und β reguläre Ausdrücke sind, dann ist $\mathcal{L}((\alpha \cup \beta)) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$.

(5) Falls α ein regulärer Ausdruck ist, dann ist $\mathcal{L}(\alpha^*) = \mathcal{L}(\alpha)^*$.

Für alle regulären Ausdrücke α, β und γ gelten:

$\alpha(\beta \cup \gamma) \equiv \alpha\beta \cup \alpha\gamma$	(Distributivität von Links)
$(\alpha \cup \beta)\gamma \equiv \alpha\gamma \cup \beta\gamma$	(Distributivität von Rechts)
$\alpha \cup (\beta \cup \gamma) \equiv (\alpha \cup \beta) \cup \gamma$	(Assoziativität für $\cup$)
$\alpha(\beta\gamma) \equiv (\alpha\beta)\gamma$	(Assoziativität für $\circ$)
$\alpha \cup \beta \equiv \beta \cup \alpha$	(Kommutativität für $\cup$)
$\varepsilon \circ \alpha \equiv \alpha \circ \varepsilon \equiv \alpha$	(Neutralität von ε bzgl. $\circ$)
$\emptyset \cup \alpha \equiv \alpha \cup \emptyset \equiv \alpha$	(Neutralität von $\emptyset$ bzgl. $\cup$)
$\emptyset \circ \alpha \equiv \alpha \circ \emptyset \equiv \emptyset$	(Dominanz von $\emptyset$ bzgl. $\circ$)

Beweis/widerlegen: ist es eine

- reguläre Sprache

- nicht reguläre Sprache
- kontextfrei

Umformung zu Chomsky Normalform
Abgeschlossenheit

Eindeutig, Mehrdeutig

Satz (Pumping Lemma):
Sei L eine reguläre Sprache. Dann gibt es eine Zahl n, so dass sich alle Wörter $w \in L$ mit $|w| \geq n$ in $w = xyz$ zerlegen lassen, so dass

(1) $y \neq \varepsilon$
(2) $|xy| \leq n$
(3) $xy^kz \in L$ *für alle* $k \geq 0$

$xz \in L$
$w = xyz \in L$
$xyyz \in L$
$xyyyz \in L$

Typ 0 *allgemeine Sprachen*

Typ 1 *kontextsensitive Sprachen*

Typ 2 *kontextfreie Sprachen*

Typ 3 *reguläre Sprachen*

Transformation einer kontextfreien Grammatik in Chomsky-Normalform:

(1) Elimination des Startsymbols auf rechten Seiten
(2) Elimination von ε-Regeln
(3) Elimination von Kettenregelzyklen
(4) Elimination von Kettenregeln
(5) Elimination von nichtisolierten Terminalsymbolen auf rechten Seiten
(6) Elimination von langen rechten Seiten

Definition:
Ein Kellerautomat ist ein 6-Tupel $(K, \Sigma, \Gamma, \Delta, s, F)$, wobei gilt:

- *K ist eine endliche Menge von Zuständen,*
- *Σ ist ein Alphabet, das Eingabealphabet,*
- *Γ ist ein Alphabet, das Kelleralphabet,*
- *$s \in K$ ist der Startzustand,*
- *$\Delta \subseteq (K \times (\Sigma \cup \{\varepsilon\}) \times \Gamma^*) \times (K \times \Gamma^*)$ ist die Übergangsrelation,*
- *Δ ist endlich und*
- *$F \subseteq K$ ist die Menge der Endzustände.*

Satz (Pumping Lemma):
Sei L eine kontextfreie Sprache. Dann gibt es eine Zahl n, so dass sich alle Wörter $z \in L$ mit $|z| \geq n$ in $z = uvwxy$ zerlegen lassen, so dass

(1) $vx \neq \varepsilon$
(2) $|vwx| \leq n$
(3) $uv^iwx^iy \in L$ *für alle $i \geq 0$*

Definition:
Eine deterministische Turing-Maschine *ist ein 7-Tupel* $(K,\Sigma,\Gamma,\delta,s,q_{\text{accept}},q_{\text{reject}})$, *wobei gilt:*

- K *ist eine endliche Menge von* Zuständen,
- Σ *ist ein Alphabet, das* Eingabealphabet, *das das* Blanksymbol $\sqcup$ *nicht enthält,*
- Γ *ist ein Alphabet, das* Bandalphabet, *das alle Elemente von* Σ *und das Blanksymbol* $\sqcup$ *enthält,*
- $s \in K$ *ist der* Startzustand,
- $q_{\text{accept}} \in K$ *ist der* akzeptierende Halteszustand,
- $q_{\text{reject}} \in K$ *ist der* verwerfende Halteszustand, $q_{\text{reject}} \neq q_{\text{accept}}$,
- $\delta : (K - \{q_{\text{accept}}, q_{\text{reject}}\}) \times \Gamma \to K \times \Gamma \times \{L,R,S\}$ *ist die* Übergangsfunktion.

Definition:
Eine deterministische Turing-Maschine *ist ein 7-Tupel* $(K,\Sigma,\Gamma,\delta,s,q_{\text{accept}},q_{\text{reject}})$, *wobei gilt:*

- K *ist eine endliche Menge von* Zuständen,
- Σ *ist ein Alphabet, das* Eingabealphabet, *das das* Blanksymbol $\sqcup$ *nicht enthält,*
- Γ *ist ein Alphabet, das* Bandalphabet, *das alle Elemente von* Σ *und das Blanksymbol* $\sqcup$ *enthält,*
- $s \in K$ *ist der* Startzustand,
- $q_{\text{accept}} \in K$ *ist der* akzeptierende Halteszustand,
- $q_{\text{reject}} \in K$ *ist der* verwerfende Halteszustand, $q_{\text{reject}} \neq q_{\text{accept}}$,
- $\delta : (K - \{q_{\text{accept}}, q_{\text{reject}}\}) \times \Gamma \to K \times \Gamma \times \{L,R,S\}$ *ist die* Übergangsfunktion.

Definition:
Eine **nichtdeterministische Turing-Maschine** *ist ein 7-Tupel* $(K, \Sigma, \Gamma, \Delta, s, q_{\text{accept}}, q_{\text{reject}})$, *wobei gilt:*

- K *ist eine endliche Menge von Zuständen,*
- Σ *ist das Eingabealphabet,* $\sqcup \notin \Sigma$,
- Γ *ist das Bandalphabet,* $\Gamma \supseteq \Sigma$ *und* $\sqcup \in \Gamma$,
- $s \in K$ *ist der Startzustand,*
- $q_{\text{accept}} \in K$ *ist der akzeptierende Halteszustand,*
- $q_{\text{reject}} \in K$ *ist der verwerfende Halteszustand,* $q_{\text{reject}} \neq q_{\text{accept}}$,
- $\Delta \subseteq (K - \{q_{\text{accept}}, q_{\text{reject}}\}) \times \Gamma \times K \times \Gamma \times \{L, R, S\}$ *ist die* **Übergangsrelation**.

Satz:
Die Klasse der entscheidbaren Sprachen ist abgeschlossen unter

(a) Vereinigung,
(b) Konkatenation,
(c) Kleene Star,
(d) Komplement und
(e) Schnitt.

Satz von Rice

Satz: [Rice]
Sei Σ *ein Alphabet und* $\mathcal{T}$ *eine echte, nicht-leere Teilmenge der Menge aller rekursiv aufzählbaren Sprachen über* Σ. *Dann ist*

$$\mathscr{L}_{\mathcal{T}} = \{\langle M \rangle \mid \mathcal{L}(M) \in \mathcal{T}\}$$

nicht entscheidbar.

DEA konstruieren
Turing machinen

reguläre Ausdrücke angeben

Knotenüberdeckung