

DEFINITION

Alphabet

DEFINITION

Menge der endlichen Folgen

DEFINITION

Wort

DEFINITION

Induktiv w^n definieren

DEFINITION

y, w sind Wörter über Σ . Dann heißt y :

DEFINITION

Sprachen

DEFINITION

Präfix

DEFINITION

Infix

DEFINITION

Suffix

DEFINITION

formale Sprachen

Für eine Menge X ist X^* die Menge der endlichen Folgen über X .

Beispiel: Elemente von $a, b, c, d^* : (a, b, c), ()$

$$w^n = \begin{cases} \epsilon & \text{falls } n = 0 \\ w * w^{n-1} & n > 0 \end{cases}$$

f: Menge der möglichen Eingaben \rightarrow Menge der möglichen Ausgaben
Spezialfall $A = 0, 1$ heißt Entscheidungsproblem. Sie ist gegeben durch die Menge der Eingaben.

Seien y, w Wörter über Σ . Dann heißt Infix/Faktor von w , wenn es $x, z \in \Sigma^*$ gibt mit $xyz = w$.

Sei Σ ein Alphabet. Teilmengen von Σ^* werden formale Sprachen über Σ genannt.
Eine Menge L ist eine formale Sprache wenn es ein Alphabet Σ gibt, so dass L formale Sprache über Σ ist (d.h. $L \subseteq \Sigma^*$).

Ein Alphabet ist eine endliche nichtleere Menge. Üblicherweise heißen Alphabete hier Σ, Γ, Δ . Ist Σ Alphabet, so nennen wir die Elemente oft Buchstaben und die Elemente von Σ^* auch Wörter über Σ (auch String/Zeichenkette).

Sind $u = (a_1, a_2, \dots, a_n)$ und $v = (b_1, b_2, \dots, b_n)$ Wörter, so ist $u * v$ das Wort $(a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n)$; es wird als Verkettung/Konkatenation von u und v bezeichnet. An Stelle von $u * v$ schreibt man auch uv .

- Präfix/Anfangsstück von w , wenn es $z \in \Sigma^*$ gibt mit $yz = w$
- Infix/Faktor von w , wenn es $x, z \in \Sigma^*$ gibt mit $xyz = w$
- Suffix/Endstück von w , wenn es $x \in \Sigma^*$ gibt mit $xy = w$

Seien y, w Wörter über Σ . Dann heißt Präfix/Anfangsstück von w , wenn es $z \in \Sigma^*$ gibt mit $yz = w$.

Seien y, w Wörter über Σ . Dann heißt Suffix/Endstück von w , wenn es $x \in \Sigma^*$ gibt mit $xy = w$.

DEFINITION

Verkettung von Sprachen

DEFINITION

Kleene Abschluss

DEFINITION

Prioritätsregeln für Operationen auf Sprachen

DEFINITION

Grammatik

DEFINITION

Ableitung einer Grammatik

DEFINITION

Wort ist Satzform

DEFINITION

erzeugte Sprache

DEFINITION

Chomsky-0

DEFINITION

Chomsky-1

DEFINITION

Chomsky-2

Sei L eine Sprache. Dann ist $L^* = \bigcup_{n \geq 0} L^n$ der Kleene-Abschluss oder die Kleene-Iteration von L .

Weiter ist $L^+ = \bigcup_{n \geq 0} L^n$
 $(L^+ = L * L = L^* * L)$

Sind L_1 und L_2 Sprachen, so heißt die Sprache $L_1 L_2 = \{w \mid \exists w_1 \in L_1, w_2 \in L_2 : w = w_1 w_2\}$ (auch $L_1 * L_2$) die Konkatenation oder Verkettung von L_1 und L_2 .

Grammatiken sind ein Mittel um alle syntaktisch korrekten Sätze einer Sprache zu erzeugen.

Eine Grammatik G ist ein 4-Tupel $G = (V, \Sigma, P, S)$ das folgende Bedingungen erfüllt

- V ist eine endliche Menge von Nicht-Terminalen oder Variablen
- Σ ist ein Alphabet (Menge der Terminale) mit $V \cap \Sigma = \emptyset$, d.h. kein Zeichen ist gleichzeitig Terminal und Nicht-Terminal
- $P \subseteq (V \cup \Sigma)^+ \times (V \cup \Sigma)^*$ ist eine endliche Menge von Regeln oder Produktionen (Produktionsmenge)
- $S \in V$ ist das Startsymbol/ die Startvariable oder das Axiom

Jede Grammatik hat nur endlich viele Regeln!

- Potenz/Iteration binden stärker als Konkatenation
- Konkatenation stärker als Vereinigung/Durchschnitt/Differenz

Ein Wort $w \in (V \cup \Sigma)^*$ heißt Satzform, wenn es eine Ableitung gibt, deren letztes Wort w ist.

Sei $G = (V, \Sigma, P, S)$ eine Grammatik. Eine Ableitung ist eine endliche Folge von Wörtern $w_0, w_1, w_2, \dots, w_n$ mit $w_0 \Rightarrow w_1 \Rightarrow w_2 \Rightarrow \dots \Rightarrow w_n$.

Jede Grammatik ist vom Typ 0 (Semi-Thue-System) und wird auch als rekursiv-aufzählbar bezeichnet.

Die Sprache $L(G) = \{w \in \Sigma^* \mid S \Rightarrow_G^* w\}$ aller Satzformen aus Σ^* heißt von G erzeugte Sprache.

Eine Regel $(l \rightarrow r)$ heißt kontext-frei wenn $l \in V$ und $r \in (V \cup \Sigma)^*$ gilt. Eine Grammatik ist vom Typ 2, falls sie nur kontext-freie Regeln enthält

Eine Regel heißt kontext-sensitiv, wenn es Wörter $u, v, w \in (V \cup \Sigma)^*$, $|v| > 0$ und ein Nichtterminal $A \in V$ gibt mit $l = uAw$ und $r = vww$. Eine Grammatik ist vom Typ 1 (oder kontext-sensitiv) falls

- alle Regeln aus P kontext-sensitiv sind
- $(S \rightarrow \epsilon) \in P$ die einzige nicht kontext-sensitive Regel in P ist und S auf keiner rechten Seite einer Regel aus P vorkommt

DEFINITION

Chomsky-3

BEWEISE

Es gibt einen Algorithmus, der als Eingabe eine Typ-1-Grammatik G und ein Wort w bekommt und nach endlicher Zeit entscheidet ob $w \in L(G)$ gilt.

DEFINITION

Deterministische endliche Automaten

DEFINITION

DFA mit Funktion $\hat{\delta}$

DEFINITION

von einem DFA akzeptierte Sprache

DEFINITION

Wann ist eine Sprache regulär?

DEFINITION

Text

1. $w = \epsilon$: Da G vom Typ 1 ist, gilt $w \in L(G)$ genau dann wenn $(S \rightarrow \epsilon) \in P$. Dies kann ein Algorithmus entscheiden

2. $|w| \geq 1$: Definiere einen gerichteten Graphen (W, E) wie folgt

- Knoten sind die nichtleeren Wörter über $V \cup \Sigma$ der Länge $\geq |w|$ (insbes. $S, w \in W$)
- $(u, v) \in E$ genau dann wenn $u \Rightarrow_G v$

da kontext-sensitiv ist, gilt $1 = |u_0| \geq |u_1| \geq |u_2| \geq \dots \geq |u_n| = |w|$, also $u_i \in W$ f.a. $1 \leq i \leq n$. Also existiert Pfad von S nach w im Graphen (W, E) , womit die Behauptung bewiesen ist.

Zu einem gegebenen DFA definieren wir die Funktion

$$\hat{\delta} : Z \times \Sigma^* \rightarrow Z \text{ induktiv wie folgt, wobei } z \in Z, \\ w \in \Sigma^+ \text{ und } a \in \Sigma:$$

- $\hat{\delta}(z, \epsilon) = z$
- $\hat{\delta}(z, aw) = \hat{\delta}(\delta(z, a), w)$

Der Zustand $\hat{\delta}(z, w)$ ergibt sich indem man vom Zustand z aus dem Pfad folgt der mit w beschriftet ist.

Eine Sprache $L \subseteq \Sigma^*$ ist regulär, wenn es einen DFA mit $L(M) = L$ gibt (bzw. wird von einem DFA akzeptiert). Jede reguläre Sprache ist rechtslinear.

Eine Regl ist rechtslinear, wenn $l \in V$ und $r \in \Sigma V \cup \epsilon$ gilt. Eine Grammatik ist vom Typ 3 wenn sie nur rechtslineare Regeln enthält.

ein deterministischer endlicher Automat M ist ein 5-Tupel $M = (Z, \Sigma, z_0, \delta, E)$

- Z eine endliche Menge von Zuständen
- Σ das Eingabealphabet (mit $Z \cap \Sigma = \emptyset$)
- $z_0 \in Z$ der Start/Anfangszustand (max Einer)
- $\delta : Z \times \Sigma \rightarrow Z$ die Überführungs/Übergangsfunktion
- $E \subseteq Z$ die Menge der Endzustände

Abkürzung: DFA (deterministic finite automaton)

die von einem DFA akzeptierte Sprache ist:

$$L(M) = \{w \in \Sigma^* \mid \hat{\delta}(z_0, w) \in E\}$$

Mit anderen Worten: Ein Wort w wird genau dann akzeptiert, wenn derjenige Pfad, der im Anfangszustand beginnt und dessen Übergänge mit den Zeichen von w markiert sind, in einem Endzustand endet.

Text