Definition

Alphabet

Definition

Menge der endlichen Folgen

Definition

Wort

Definition

Induktiv w^n definieren

Definition

y,w sind Wörter über $\sum$. 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 $a, b, c, d* : (a, b, c), ()$

Ein Alphabet ist eine endliche nichtleere Menge. Üblicherweise heißen Alphabete hier $\sum, \Gamma, \Delta$. Ist $\sum$ Alphabet, so nennen wir die Elemente oft Buchstaben und die Elemente von $\sum *$ auch Wörter über $\sum$ (auch String/Zeichenkette).

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

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

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

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

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

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

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

Seien y,w Wörter über $\sum$. Dann heißt Suffix/Endstück von w, wenn es $x \in \sum *$ 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 $L_1L_2 = \{w|\exists w_1 \in L_1, w_2 \in L_2 : w = w_1w_2\}$ (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 $G = (V, \sum, P, S)$ das folgende Bedingungen erfüllt

- V ist eine endliche Menge von Nicht-Terminalen oder Variablen
- $\sum$ ist ein Alphabet (Menge der Terminale) mit $V \cap \sum = \varnothing$,d.h. kein Zeichen ist gleichzeitig Terminal und Nicht-Terminal
- $P \subseteq (V \cup \sum)^+ \times (v \cup \sum)^*$ 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 $w \in (V \cup \sum)^*$ heißt Satzform, wenn es eine Ableitung gibt, deren letztes Wort w ist.

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

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

Die Sprache $L(G) = w \in \sum^* |S \Rightarrow_G^* w$ aller Satzformen aus $\sum^*$ heißt von G erzeugte Sprache.

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

Eine Regel heißt kontext-sensitiv, wenn es Wörter $u, v, w \in (V \cup \sum)^*, |v| > 0$ und ein Nichtterminal $A \in V$ gibt mit $l = uAw$ und $r = uvw$. 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 bekommst und nach endlicher Zeit entscheidet ob $w \in L(G)$ gilt.

Definition

Deterministische endliche Automaten

Definition

DFA mit Funktion $\hat{\hat{\delta}}$

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 $(S \to \epsilon) \in P$. Dies kannn ein Algorithmus entscheiden

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

   - Knoten sind die nichtleeren Wörter über $V \cup \sum$ 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 $1 = |u_0| \geq |u_1| \geq |u_2| \geq ... \geq |u_n| = |w|$, also $u_i \in W$ f.a. $1 \geq i \geq n$. Also existiert Pfad von S nach w im Graphen (W , E), womit die Behauptung bewiesen ist.

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

Zu einem gegebenen DFA definieren wir die Funktion $\hat{\delta} : Z \times \sum^* \to Z$ induktiv wie folgt, wobei $z \in Z$, $w \in \sum^+$ und $a \in \sum$:

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

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

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

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

Abkürzung: DFA (deterministic finite automaton)

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

die von einem DFA akzeptierte Sprache ist: $L(M) = w \in \sum^* | \hat{\delta}(z_0, w) \in E$
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