

endliche Automaten (Maschinen)

Reguläre Ausdrücke

Nicht-Reguläre Sprachen

Entscheidbarkeit

DFA M

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

Eine Sprache $L \supseteq \sum^*$ ist regulär, wenn es einen DFA mit L(M) = L gibt

NFA M

Jede von einem NFA akzeptierte Sprache ist regulär

Satz

Wenn L_1 und L_2 reguläre Sprachen sind, dann ist auch

 $L_1 \cup L_2$ regulär

 $L_1 \cap L_2$ regulär

 L_1L_2 regulär

 L_1^+/L_1^* regulär

Jede reguläre Sprache ist rechtslinear

Definition

Die Menge $Reg(\sum)$ der regulären Ausdrücke über dem Alphabet \sum ist die kleinste Menge mit folgenden Eigenschaften:

 $\varnothing \in Reg(\sum), \lambda \in Reg(\sum), \sum \subseteq Reg(\sum)$

Wenn $\alpha, \beta \in Reg(\sum)$, dann auch $(\alpha * \beta), (\alpha + \beta), (\alpha^*) \in Reg(\sum)$

für $\alpha * \beta$ schreibt man oft $\alpha\beta$

für $\alpha + \beta$ schreibt man auch $\alpha | \beta$

Für einen regulären Ausdruck $\alpha \in Reg(\sum)$ ist die Sprache $L(\alpha) \subseteq \sum^*$ induktiv definiert

zu jedem regulären Ausdruck γ gibt es einen NFA M mit $L(\gamma) = L(M)$

zu jedem DFA M gibt es einen regulären Ausdruck γ mit $L(M) = L(\gamma)$

Pumping Lemma

L sei reguläre Sprache, dann gibt es $n \leq 1$ derart, dass für alle $x \in L$ mit $|x| \geq n$ gilt: es gibt Wörter $u, v, w \in \sum^*$ mit

 $x = uvw, |uv| \le n, |v| \ge 1$

 $uv^iw \in L$ für alle $i \geq 0$

geeignet um Aussagen über Nicht-Regularität zu machen

Myhill-Nerode Äquivalenz

binäre Relation $R_L \subseteq \sum^* \times \sum^*$

 $\forall x, y \in \sum^* \text{ setze } (x, y) \in R_L$ genau dann, wenn $\forall z \in \sum^* :$ $(xy \in L \leftrightarrow yz \in L) \text{ gilt. } xR_Ly$

Für Sprache L und Wort $x \in \sum^*$ ist $[x]_L = \{y \in \sum^* | xR_L y\}$ die Äquivalenzklasse von x

Satz: L ist regulär $\leftrightarrow index(R_L) < \infty$

Wortproblem

Gilt $w \in L$ für eine gegebene reguläre Sprache L und $w \in \sum^*$?

Leerheitsproblem

Gilt $L = \emptyset$ für eine gegebene reguläre Sprache L?

Endlichkeitsproblem

Ist eine gegebene reguläre Sprache L endlich?

Schnittproblem

Gilt $L_1 \cap L_2 = \emptyset$ für gegebene reguläre L_1, L_2 ?

Inklusionsproblem

Gilt $L_1 \subseteq L_2$ für gegebene reguläre L_1, L_2 ?

Äquivalenzproblem

Gilt $L_1 = L_2$ für gegebene reguläre L_1, L_2 ?

deterministischer endlicher Automat M

- 5-Tupel $M = (Z, \sum, z_0, \delta, E)$
- Z eine endliche Menge von Zuständen
- \sum das Eingabealphabet (mit $Z \cap \sum = \emptyset$)
- $\overline{z_0} \in Z$ der Startzustand
- $\delta: Z \times \sum \to Z$ die Übergangsfunktion $E \subseteq Z$ die Menge der Endzustände
- kurz: DFA (deterministic finite automaton)

Turingmaschine

Definition: Eine Turingmaschine (TM) ist ein 7-Tupel

$$M = (Z, \sum, \Phi, \delta, z_o, \Box, E)$$

, wobei

- \sum das Eingabealphabet
- Φ mit $\Phi \supseteq \sum$ und $\Phi \cap Z \neq 0$ das Arbeits- oder Ban-
- $z_0 \in Z$ der Startzustand,
- δ : $Z \times \Phi \rightarrow (Z \times \Phi \times \{L, N, R\})$ die bei Eingabe von x hält. Uberführungsfunktion
- $\Box \in \Phi / \Sigma$ das Leerzeichen oder Blank und
- $E \subseteq Z$ die Menge der Endzustände ist

Linksableitung

Eine Ableitung $S = w_0 \Rightarrow w_1 \Rightarrow ... \Rightarrow w_r = w \in \sum^*$ heißt Linksableitung, wenn in jedem Schritt das am weitesten links stehende Nichtterminal ersetzt wird (Analog Rechtsableitung).

CYK-Algorithmus

Gehört ein gegebenes Wort zu L(G)?

	a_1	a_2	 	a_{n-1}	an
j = 1	$V_{1,1}$	V _{2,1}	 	$V_{n-1,1}$	$V_{n,1}$
j = 2	V _{1,2}	V _{2,2}	 	$V_{n-1,2}$	
:				,	
j = n - 1	$V_{1,n-1}$	$V_{2,n-1}$			
j = n	$V_{1,n}$				

Von oben nach unten, von links unten nach rechts oben

Kellerautomaten

Um ein Automatenmodell für Kontextfreie Sprachen zu erhalten führt man einen Keller-(Pushdown)-Speicher ein, auf dem sich eine beliebig lange Sequenz von Zeichen befinden darf. Beim Einlesen eines neuen Zeichens wird das oberste Zeichen des Kellers gelesen und durch eine (evtl. leere) Sequenz von Zeichen ersetzt. An anderen Stellen kann der Keller nicht gelesen/geändert werden

die Greibach-Normalform

Eine kontextfreie Grammatik G ist in Greibach Normalform falls alle Produktionen aus P folgende Form haben: $A \to aB_1B_2...B_k$, mit $k \in \mathbb{N}, A, B_1, ..., B_k \in V$ und $a \in \sum A$ Die Greibach Normalform garantiert, dass bei jedem Ableitungsschritt genau ein Alphabetsymbol entsteht.

das Lemma von Ogden (William Ogden)

Wenn L eine kontextfreie Sprache ist, dann gibt es $n \geq 1$ derart, dass für alle $z \in L$, in denen n Positionen markiert sind, gilt: es gibt Wörter $u, v, w, x, y \in \sum^*$ mit

- 1. z = uvwxy
- 2. v oder x enthält wenigstens eine der Markierungen
- 3. $uv^iwx^iy \in L$ für alle i > 0

Halteproblem

allgemein: Das Halteproblem ist die Menge aller Paare (M,x), wobei M eine TM ist und $x \in \{0,1\}^*$, sodass M

Reduktion

Seien $A \subseteq \sum^*, B \subseteq \Phi^*$. Eine Reduktion von A auf B ist eine totale und berechenbare Funktion $f: \sum^* \to \Phi^*$, so dass für alle $w \in \sum^*$ gilt: $w \in A \leftrightarrow f(x) \in \overline{B}$. A heißt auf B reduzierbar (in Zeichen A < B), falls es eine Reduktion von A auf B gibt.

Satz von Rice

Sei R die Klasse aller Turing-berechenbaren Funktionen $\{0,1\}^* \to \{0,1\}^*$, Ω die nirgendwo definierte Funktion und sei $S \subseteq \mathbb{R}$ mit $\Omega \in S$ und $\neq \mathbb{R}$. Dann ist die Sprache $C(S) = \{w \in L_{TM} | \phi_w \in S\}$ unentscheidbar.

Semi Entscheidbarkeit

Auch wenn das Halteproblem bei leerer Eingabe H_0 unentscheidbar ist, so kann doch nach endlicher Zeit festgestellt werden, daß die Maschine M_w bei leerer Eingabe anhält -

 H_0 ist also "halb-öder ßemi-entscheidbar". Eine Sprache $L\subseteq \sum^*$ heißt semi-entscheidbar, falls die "halbe"charakteristische Funktion von L, d.h. die partielle Funktion X'_L : $\sum^* \rightarrow \{1\}$ mit x'_L = $\int 1$ falls $w \in L$ berechenbar ist. $\{undef. \text{ falls } w \notin L\}$

Universelle Turing Maschine

eine Turing-Maschine, die jede Turing-Maschine simulieren kann, wenn deren Kodierung gegeben ist. Buchstaben des Bandalphabets als Wörter über $\{0, 1, 2\}$ mit $\square = 2$ kodiert. Ab jetzt nehmen wir an, daß wir immer dieses Bandalpha-

Eine Turing Maschine U heißt universelle Turing Maschine, wenn sie die folgende partielle Funktion berechnet.

$$\{0,1\}^* \to \{0,1\}^*$$

 $y \to \begin{cases} \phi_w(x) & \text{falls } y = w000x, w \in L_{TM}, x \in \{0,1\}^* \\ undef. & \text{sonst} \end{cases}$

Totale berechenbare Funktionen

Gesucht ist $C \subseteq TOT \subseteq L_{TM}$, so dass $\{\phi_w | w \in C\} =$ $\{\phi_w|w\in TOT\}$ die Menge der totalen berechenbaren Funktionen ist.

- TOT ist nicht einmal semi-entscheidbar
 indirekt nehmen wir an, dass C semi-entscheidbar ist
- \bullet dann existiert eine totale berechenbare Funktion f: $\{0,1\}^* \to \{0,1\}^*$ mit Bild C
- neue Funktion $g: \{0,1\}^* \to \{0,1\}^* : w \vdash 1\phi_{f(w)}(w)$

Chomsky Normalform

Eine kontextfreie Grammatik G ist in Chomsky Normalform, falls alle Produktionen

- die Form $A \to AB$ oder $A \to a$ haben,
- $\bullet\,$ und $S \to \epsilon$ und S nie auf der rechten Seite einer Produktion vorkommt