Grammatiken und die Chomsky-Hierarchie

Def.: Eine Grammatik $G=(\Sigma,V,S,R)$ besteht aus

- endlichem Alphabet Σ
- endlicher Variablenmenge V mit $V\cap\Sigma=\emptyset$
- Startsymbol $S\in V$
- endlicher Menge $R \subset (V\cup\Sigma)^{+}\times(V\cup\Sigma)^{*}$ von Ableitungsregeln *("Produktionen")*

Überführungsrelation einer Grammatik

Def.: Sei $G=(\Sigma,V,S,R)$ Grammatik

- $w \in (V\cup\Sigma)^+$ wird durch G in $z \in (V\cup\Sigma)^*$ überführt, falls
  $w = w_1 l w_2$ und $z = w_1 r w_2$ und $(l,r)\in R$

- Überführungsrelation wird mit $\rightarrow_G$ bezeichnet: $w\rightarrow_G z$

- transitiver Abschluss von $\rightarrow_G$ wird mit $\rightarrow^*_G$ bezeichnet:

  $$w \rightarrow^*_G z \Leftrightarrow \exists\ w=w_1,w_2,\ldots,w_n=z\ \forall 1\leq i<n : w_{i-1}\rightarrow_G w_i$$

Von einer Grammatik erzeugte Sprache

Def.: Sei $G=(\Sigma,V,S,R)$ Grammatik.

$$L(G) := \{ w \in \Sigma^* \mid S \rightarrow^*_G w\}$$

ist die von G erzeugte Sprache

Von einer Grammatik erzeugte Sprache

Bsp.:

$\Sigma = \{(\ ,)\ , a, +, *\}$

$V = \{S\}$

$R = \{S \rightarrow (S)+(S),$

$S \rightarrow (S)*(S),$

$S \rightarrow a,$

$S \rightarrow a+a,$

$S \rightarrow a*a\ \}$

$L(G)$ = korrekt geklammerte arithmetische Ausdrücke

Von einer Grammatik erzeugte Sprache

Bsp.:

$\Sigma = \{0, 1\}$

$V = \{S\}$

$R = \{S \to 0S0,$

$S \to 1S1,$

$S \to 0,$

$S \to 1,$

$S \to \varepsilon\}$

$L(G)$ = Palindrome über $\{0,1\}$

Wortproblem einer Grammatik

Def.: Das Problem, zu einer Grammatik G zu entscheiden, ob für $w \in \Sigma^*$ gilt $w \in L(G)$, heisst Wortproblem von G

Chomsky-Hierarchie

Def.: Eine Grammatik $G=(\Sigma,V,S,R)$ heisst

- Typ-3 Grammatik ("*rechtslinear*"), falls alle Regeln folgende Form haben:

  $l \rightarrow r$ mit $l\in V$ und $r=\varepsilon$ oder $r=aB$ wobei $a\in\Sigma$, $B\in V$

- Typ-2 Grammatik ("*kontextfrei*"), falls alle Regeln folgende Form haben:

  $l \rightarrow r$ mit $l\in V$ und $r\in(V\cup\Sigma)^*$

- Typ-1 Grammatik ("*kontextsensitiv*" bzw. "*monoton*"), falls alle Regeln folgende Form haben:

  $S \rightarrow \varepsilon$ *oder* $l \rightarrow r$ mit $l\in V^+$, $r\in(V\cup\Sigma\setminus\{S\})^+$ und $|l|\leq|r|$

- Typ-0 Grammatik sonst

Chomsky-Hierarchie

Def.: Eine Sprache L heisst Typ-i Sprache, falls es eine Typ-i Grammatik G gibt, so dass $L=L(G)$

Charakterisierung von Typ-0 Sprachen

Satz: L ist Typ-0 Sprache gdw. L ist rekursiv aufzählbar

Bew.:

- L ist Typ-0 Sprache

→ es gibt Typ-0 Grammatik $G=(\Sigma,V,S,R)$ mit $L=L(G)$

Ziel: Konstruiere ntm M mit $L=L(M)$

(Arbeitsband speichert aktuelles Wort (anfangs S))

  - wähle nichtdeterministisch eine anwendbare Regel $r \in R$
  - wende r auf das aktuelle Wort an
  - vergleiche aktuelles Wort mit Eingabe w und akzeptiere w bei Übereinstimmung

Charakterisierung von Typ-0 Sprachen

Satz: L ist Typ-0 Sprache gdw. L ist rekursiv aufzählbar

Bew.:

- L ist rekursiv aufzählbar

→ es gibt dtm $M=(Q,\Sigma,\Gamma,\delta,q_0,\underline{b},F)$ mit $L=L(M)$

Ziel: KonstruiereTyp-0 Grammatik $G=(\Sigma,V,S,R)$ mit $L=L(G)$

Idee: G beschreibt Rückwärtsrechnung von (*eindeutiger*) akzeptierender Endkonfiguration zu (*eindeutiger*) Anfangskonfiguration

Charakterisierung von Typ-0 Sprachen

Bew.:

o.B.d.A. gelte für M:

- q_0 wird *nur in der Anfangskonfiguration* benutzt
- *genau ein* akzeptierender Zustand q^* existiert
- bei Akzeptanz stehen *nur* noch $\underline{b}$s auf dem Band

→ $q_0 w \underline{b}$ ist Anfangskonfiguration bei Eingabe w

→ akzeptierende Endonfigurationen sind von der Form $\underline{b}\ldots\underline{b}q^*\underline{b}\ldots\underline{b}$

Charakterisierung von Typ-0 Sprachen

Bew.:

Regeln von G:

- Anfangsregeln:

  $S \rightarrow q^*$, $q^* \rightarrow q^*\underline{b}$, $q^* \rightarrow \underline{b}q^*$

  Idee: Anfangsregeln erzeugen die *akzeptierende Endonfiguration*: $S \rightarrow^* \underline{b}\ldots\underline{b}q^*\underline{b}\ldots\underline{b}$

- Abschlussregeln: für $a \in \Gamma$

  $q_0 a \rightarrow a q_0$, $q_0\underline{b} \rightarrow \varepsilon$

  Idee: Abschlussregeln ermöglichen die Ableitung von w aus der *Anfangskonfiguration*: $q_0 w\underline{b} \rightarrow^* w$

Charakterisierung von Typ-0 Sprachen

Bew.:

Regeln von G:

- (Rückwärts)rechenregeln:
  - falls $\delta(q,a) = (q',a',R)$
    $(a'q' \rightarrow qa) \in R$
  - falls $\delta(q,a) = (q',a',S)$
    $(q'a' \rightarrow qa) \in R$
  - falls $\delta(q,a) = (q',a',L)$
    $\forall b \in \Gamma : (q'ba' \rightarrow bqa) \in R$

Offensichtlich: $L(G)=L(M)$

Wortproblem für Typ-0 Sprachen

Satz: Das Wortproblem für Typ-0 Grammatiken ist

unentscheidbar

Charakterisierung von Typ-3 Sprachen

Satz: L ist Typ-3 Sprache gdw. L ist regulär

Bew.:

- L ist Typ-3 Sprache

→ es gibt Typ-3 Grammatik $G=(\Sigma,V,S,R)$ mit $L=L(G)$

Ziel: Konstruiere nea $A=(Q,\Sigma,\delta,q_0,F)$ mit $L=L(A)$

- $Q := V$
- $q_0 := S$
- $F := \{A \in V \mid (A \rightarrow \varepsilon) \in R\}$
- $\delta(A,a) := \{B \mid (A \rightarrow aB) \in R\}$

Charakterisierung von Typ-3 Sprachen

Satz: L ist Typ-3 Sprache gdw. L ist regulär

Bew.:

- $w = w_1 \ldots w_n \in L(G)$

  → es gibt Ableitung

  $S \rightarrow w_1 A_1 \rightarrow w_1 w_2 A_2 \rightarrow \ldots \rightarrow w_1 \ldots w_n A_n \rightarrow w_1 .. w_n$

  → A hat Berechnungspfad $q_0, q_1, \ldots, q_n$, wobei

  $q_i \in \delta(q_{i-1}, w_i)$ und $q_n \in F$

  → $w \in L(A)$

- analog: $q_0 \rightarrow^* q_n \in F \rightarrow w \in L(G)$

Charakterisierung von Typ-3 Sprachen

Satz: L ist Typ-3 Sprache gdw. L ist regulär

Bew.:

- L ist regulär

→ es gibt dea $A=(Q,\Sigma,\delta,q_0,F)$ mit $L=L(A)$

Ziel: KonstruiereTyp-3 Grammatik $G=(\Sigma,V,S,R)$ mit $L=L(G)$

  - $V := Q$
  - $S := q_0$
  - $(q \rightarrow \varepsilon) \in R$ für $q \in F$
  - $(q \rightarrow aq') \in R$ falls $\delta(q,a) = q'$

Charakterisierung von Typ-3 Sprachen

Satz: L ist Typ-3 Sprache gdw. L ist regulär

Bew.:

- $w=w_1\ldots w_n \in L(A)$

→ A durchläuft Zustände $q_0,q_1,\ldots,q_n$, wobei

$$\delta(q_{i-1},w_i) = q_i \text{ und } q_n \in F$$

→ es gibt Ableitung

$$q_0 \rightarrow w_1q_1 \rightarrow w_1w_2q_2 \rightarrow \ldots \rightarrow w_1\ldots w_nq_n \rightarrow w_1\ldots w_n$$

→ $w \in L(G)$

- analog: $q_0 \rightarrow_G^* w \rightarrow w \in L(A)$

Wortproblem für Typ-3 Sprachen

Satz: Das Wortproblem für Typ-3 Grammatiken ist

entscheidbar