Chomsky-Grammatiken (Typ 0)

$$G = (\; N \;,\; T \;,\; P \;,\; S \;)$$

- N: nichtterminale Zeichen
- T: terminale Zeichen

  mit $N \cap T = \emptyset$
- P: Produktionen der Form $u ::= v$

  mit $u, v \in (N \cup T)^*$, $u \notin T^*$ (endlich viele)
- S: Startsymbol aus N

- monoton: $\mathrm{length}(u) \leq \mathrm{Length}(v)$ (Typ 1)
- kontextfrei: $\mathrm{Length}(u) = 1$ (Typ 2)
- regulär, rechtslinear: $\mathrm{Length}(u) = 1$ & $v \in T^* \cup T^+ N$ (Typ 3)

erzeugte Sprachen

- direkte Ableitung

  $$xuy \xrightarrow[u ::= v]{} xvy \quad \text{für alle } x, y \in (N \cup T)^*$$

  (1) suche u als Teilwort eines Wortes

  (2) ersetze u durch v

- Ableitung $w = w_0 \xrightarrow[r_1]{} w_1 \xrightarrow[r_2]{} w_2 \xrightarrow[r_3]{} \cdots \xrightarrow[r_n]{} w_n = w'$

  dafür kurz auch: $w \xrightarrow[P]{n} w'$ oder $w \xrightarrow[P]{*} w'$

  falls $r_1, \dots, r_n \in P$

- erzeugte Sprache

  $$L(G) = \{ w \in T^* \mid S \xrightarrow[P]{*} w \}$$

Erzeugung quadratisch langer Wörter

$G_{square} = (\{S, A, B, C, D, E, F\}, \{a\}, P_{square}, S)$ mit den Produktionen

$S ::= ABCD$, $BC ::= aCB$, $BD ::= ECCD$, $Aa ::= aA$, $Ca ::= aC$, $CE ::= EC$,
$AE ::= AB$, $ABC ::= F$, $FC ::= F$, $FD ::= \lambda$

$$S \to ABCD \to AaCBD \to aACBD \to aACECCD \to aAEC^3D \to aABC^3D \to$$

$$ABCD \to FD \to \lambda$$

$$aABC^3D \to aFC^2D \xrightarrow{2} aFD \to a$$

$$\to aAaCBC^2D \xrightarrow{2} aA(aC)^3BD \xrightarrow{5} a^4AC^3BD \to a^4AC^3EC^2D \xrightarrow{3} a^4AEC^5D \to a^4ABC^5D \to$$

$$a^4ABC^5D \to a^4FC^4D \xrightarrow{4} a^4FD \to a^4$$

$$\to \dots \to a^{n^2}ABC^{2n+1}D \xrightarrow{*} a^{(n+1)^2}ABC^{2(n+1)+1}D \xrightarrow{*}$$

$$a^{n^2}ABC^{2n+1}D \xrightarrow{*} a^{n^2}$$

$$a^{(n+1)^2}ABC^{2(n+1)+1}D \xrightarrow{*} a^{(n+1)^2}$$

Wortprobleme

- Wortproblem von $L \subseteq T^*$: für alle $w \in T^*$ als Eingaben ist zu entscheiden, ob $w \in L$?
- Wortproblem $WP(G)$ einer Grammatik G ist Wortproblem von $L(G)$
- $WP(G) \in O(n^3)$ für kontextfreier G nach Cocke-Kasami-Younger
- $WP(G) \in O(n)$, falls deterministischer Kellerautomat K existiert mit $L(G) = L(K)$
- insbes. $WP(G) \in O(n)$ für rechtslineares G wegen endlicher Automaten

Wortproblem monotoner Grammatiken in PSPACE

Beobachtung: $WP(G) \in NPSPACE = PSPACE$ für monotones G

Beweis: Sei $G = (N, T, P, S)$ monotone Grammatik und $w_0 \in T^*$.
Betrachte folgenden nichtdeterministischen Algorithmus:
Konfigurationen: (w, w') mit $Length(w') \leq Length(w)$
Berechnungsschritt: $(w, w') \vdash (w, w'')$ falls $w' \xrightarrow[P]{} w''$
nichtdeterministisch
Start: (w_0, S)
Ziel: (w_0, w_0)

Platzbedarf: $2 \cdot Length(w_0)$

Korrektheit: $w_0 \in L(G)$ gdw. $(w_0, S) \vdash^* (w_0, w_0)$

Lemma: $w_1 \xrightarrow[P]{*} w_2$ impliziert $Length(w_1) \leq Length(w_2)$

Chomsky-Hierarchie (Grammatiken & Sprachen)

Typ	Bez.	Automaten	Wortproblem		
0	allg.	Turing-Maschinen	(Reduktion Haltepr.)		
1	kontext-sensitiv, monoton	Linear-beschränkte Automaten	PSPACE = NPSPACE		
2	kontext-frei	Keller-automaten	$O(n^3)$ (CKY)		
3	rechts-linear, regulär	endliche Automaten (auch det.)	$O(n)$ (←)		

Wortproblem für monotone Grammatiken

Sei $G = (N, T, P, S)$ eine monotone Grammatik und $w_0 \in T^*$.
Dann kann man folgendermaßen algorithmisch feststellen, ob $w_0 \in L(G)$. Dafür sei $n = \text{Length}(w_0)$.

(i) $S_0 = \{S\}$,

(ii) $S_{k+1} = S_k \cup \{w \in (N \cup T)^* \mid v \xrightarrow[P]{} w,\ v \in S_k,\ \text{Length}(w) \leq n\}$,

(iii) wähle kleinstes m mit $S_m = S_{m+1}$.

Dann gilt: $S_m = \{w \in (N \cup T)^* \mid S \xrightarrow[P]{*} w,\ \text{Length}(w) \leq n\}$

Und damit insbesondere: $w_0 \in L(G)$ g.d.w. $w_0 \in S_m$.

Beweis

(1) Existenz von m: $S_0 \subseteq S_1 \subseteq \cdots \subseteq S_k \subseteq S_{k+1} \subseteq \cdots \subseteq K_n = \{w \in (N \cup T)^* \mid \mathrm{Length}(w) \leq n\}$.

endlich

(2) $w \in S_k$ impliziert $S \xrightarrow[p]{*} w$.

(IA) $k=0$: $w \in S_0 = \{S\}$ impliziert $w = S$. Es gilt $S \xrightarrow[p]{0} S = w$.

(IS) $k+1$: $w \in S_{k+1}$ impliziert $w \in S_k$ und damit nach (IV) $S \xrightarrow[p]{*} w$ oder $v \xrightarrow[p]{} w$ für $v \in S_k$. Dann gilt $S \xrightarrow[p]{*} v$ nach (IV) und somit $S \xrightarrow[p]{*} v \xrightarrow[p]{} w$.

(3) $S \xrightarrow[p]{\ell} w$ mit $\mathrm{Length}(w) \leq n$ impliziert $w \in S_m$.

(IA) $S \xrightarrow[p]{0} w$ impliziert $w = S \in \{S\} = S_0 \subseteq S_m$.

(IS) $S \xrightarrow[p]{\ell+1} w$ impliziert $S \xrightarrow[p]{\ell} v$ und $v \xrightarrow[p]{} w$.
Wegen der Monotonie gilt $\mathrm{Length}(v) \leq n$,
so dass nach (IV) folgt $v \in S_m$.
Somit $w \in S_{m+1} = S_m$.

Chomsky-Hierarchie (Grammatiken & Sprachen)

Typ	Bez.	Automaten	Wortproblem	Leerheitspr.	
0	allg.	Turing-Maschinen	$\overline{\text{(Reduktion}}$ Haltepr.)	−	
1	kontext-sensitiv, monoton	Linear beschränkte Automaten	PSPACE = NPSPACE	−	
2	kontextfrei	Kellerautomaten	$O(n^3)$ (CKY)	+	
3	rechtslinear, regulär	endliche Automaten (auch det.)	$O(n)$ (←)	+	

Leerheitsproblem für kontextfreie Grammatiken

Sei $G = (N, T, P, S)$ eine kontextfreie Grammatik.
Dann lässt sich $N_{terminate} = \{A \in N \mid A \overset{*}{\underset{P}{\Rightarrow}} u, u \in T^*\}$ folgendermaßen algorithmisch bestimmen:

(i) $N_0 = \{A \in N \mid (A ::= u) \in P, u \in T^*\}$,

(ii) $N_{k+1} = N_k \cup \{A \in N \mid (A ::= u) \in P, u \in (N_k \cup T)^*\}$,

(iii) wähle kleinstes m mit $N_m = N_{m+1}$.

Dann gilt: $N_{terminate} = N_m$.

Chomsky-Hierarchie (Grammatiken & Sprachen)

Typ	Bez.	Automaten	Wortproblem	Leerheitspr.	Durchschnitts-leerheitspr.
0	allg.	Turing-Maschinen	– (Reduktion Haltepr.)	–	–
1	kontext-sensitiv monoton	Linear beschränkte Automaten	PSPACE = NPSPACE	–	–
2	kontext-frei	Keller-automaten	$O(n^3)$ (CKY)	+	–
3	rechts-linear, regulär	endliche Automaten (auch det.)	$O(n)$ ($\leftarrow$)	+	+

Postsches Korrespondenzproblem

- PCP $= ((u_1, \ldots, u_n), (v_1, \ldots, v_n))$ mit $u_i, v_j \in T^*$
- PCP lösbar, wenn $i_1 \cdots i_k$ mit $k \leq 1$ existiert, derart dass

$$u_{i_1} \cdots u_{i_k} = v_{i_1} \cdots v_{i_k}.$$

Theorem: Lösbarkeit Postscher Korrespondenzprobleme ist nicht entscheidbar (für $\emptyset \neq T \neq \{a\}$).

Beispiel: $((aabb, a, aba), (b, aa, bab))$

Lösungsbeispiel: 2 2 1 3 2

mit $a \mid a \mid aabb \mid aba \mid a = aa \mid aa \mid b \mid bab \mid aa$

Durchschnittsleerheitsproblem für kontextfreie Grammatiken

▷ G_{mirror} mit den Produktionen $S ::= \$ \mid xSx$ für $x \in T$.

Es gilt: $L(G_{mirror}) = \{ w \$ \, trans(w) \mid w \in T^* \}$.

▷ G_{PCP} mit den Produktionen $\left.\begin{array}{l} S ::= u_i A \, trans \, v_i \\ A ::= \$ \mid u_i A \, trans \, v_i \end{array}\right\} i = 1, \ldots, n$

Es gilt: $L(G_{PCP}) = \{ u_{i_1} \cdots u_{i_k} \$ \, trans(v_{i_1} \cdots v_{i_k}) \mid 1 \leq i_j \leq n, \ k \geq 1 \}$

Beobachtung: PCP lösbar g.d.w. $L(G_{PCP}) \cap L(G_{mirror}) \neq \emptyset$.

Theorem: Das Durchschnittsleerheitsproblem für kontextfreie Grammatiken ist nicht entscheidbar.

dieses ist aus meiner Sicht das beste Buch über theoretische Informatik