Entscheidbarkeit

Wir geben Algorithmen an, mit denen übliche Probleme für kontextfreie Sprachen gelöst werden können.

Wortproblem für eine kontextfreie Sprache L

Gegeben $w \in \Sigma^*$.
Gilt $w \in L$?

Ist die kontextfreie Sprache L durch eine kontextfreie Grammatik in Chomsky-Normalform gegeben, so kann das Wortproblem mit dem CYK-Algorithmus in Zeit $O(|w|^3)$ gelöst werden.

Ist L durch einen deterministischen PDA gegeben, so kann das Wortproblem für L sogar in Zeit $O(n)$ gelöst werden.

Uniformes Wortproblem für kontextfreie Sprachen

Gegeben: kontextfreie Grammatik G und Wort $w \in \Sigma^*$.
Gilt $w \in L(G)$?

Lösung:

1. Berechne kontextfreie Grammatik G' in Chomsky-Normalform mit $L(G) = L(G')$.
2. Wende CYK-Algorithmus auf die Frage „$w \in L(G')$?“ an.

Leerheitsproblem für kontextfreie Sprachen

Gegeben eine kontextfreie Grammatik $G = (V, \Sigma, P, S)$.
Gilt $L(G) = \emptyset$?

Lösung:
Sei

$$W = \{A \in V \mid \exists w \in \Sigma^* \colon A \Rightarrow_G^* w\}$$

die Menge aller produktiven Nichtterminale. Dann gilt
$L(G) \neq \emptyset \iff S \in W$.

Berechnung von W:

$W_0 := \{A \in V \mid \exists w \in \Sigma^* \colon (A \to w) \in P\}$
for $i = 0$ **to** $|V|$ **do**
  $W_{i+1} := W_i \cup \{A \in V \mid \exists v \in (\Sigma \cup W_i)^* \colon (A \to v) \in P\}$
endfor

Beispiel:

Sei G die kontextfreie Grammatik mit den folgenden Produktionen:

$$\begin{array}{lcl} S & \rightarrow & AC \\ A & \rightarrow & BC \\ B & \rightarrow & CA \mid b \\ C & \rightarrow & a \end{array}$$

Wir haben

1. $W_0 = \{B, C\}$,
2. $W_1 = \{A, B, C\}$ und
3. $W_2 = W_3 = W_4 = W_5 = \{S, A, B, C\}$.

Also sind alle Nichtterminale produktiv. Insbesondere gilt $L(G) \neq \emptyset$.

Behauptung

$W_{|V|}$ ist die Menge der produktiven Nichtterminale W.

Beweis:

1. Zunächst zeigen wir $W_i \subseteq W$ per Induktion über i:

   IA $W_0 \subseteq W$ ist offensichtlich

   IS Sei $A \in W_{i+1} \setminus W_i$. Dann existiert $(A \to v) \in P$ mit $v \in (\Sigma \cup W_i)^*$. Nach der IV folgt $v \in (\Sigma \cup W)^*$. Also existiert $w \in \Sigma^*$ mit $v \Rightarrow_G^* w$ und damit $A \Rightarrow_G^+ w$, d.h. $A \in W$.

2. Jetzt zeigen wir $W \subseteq W_{|V|}$:

Vorbemerkung:
Wir haben $W_0 \subseteq W_1 \subseteq W_2 \cdots \subseteq W_{|V|} \subseteq W_{|V|+1} \subseteq V$. Da V endlich ist, gibt es $k \leq |V|$ mit $W_k = W_{k+1}$. Also gilt $W_{|V|} = W_{|V|+1}$.

Sei $A \in W$. Dann existieren Wörter $x_i \in (\Sigma \cup V)^*$ mit

$$A = x_\ell \Rightarrow_G x_{\ell-1} \Rightarrow_G x_{\ell-2} \Rightarrow_G \cdots \Rightarrow_G x_0 \in \Sigma^* .$$

Per Induktion über i zeigen wir $x_i \in (\Sigma \cup W_{|V|})^*$ (also insbes. $A \in W_{|V|}$):

IA $x_0 \in \Sigma^* \subseteq (\Sigma \cup W_{|V|})^*$.

IS Es gibt $u, v, w \in (\Sigma \cup V)^*$ und $(B \to v) \in P$ mit $x_{i+1} = uBw$ und $uvw = x_i \in (\Sigma \cup W_{|V|})^*$. Damit gelten $u, v, w \in (\Sigma \cup W_{|V|})^*$ und $B \in W_{|V|+1} = W_{|V|}$, also $x_{i+1} \in (\Sigma \cup W_{|V|})^*$. $\square$

Endlichkeitsproblem für kontextfreie Sprachen

Gegeben eine kontextfreie Grammatik $G = (V, \Sigma, P, S)$.
Ist $L(G)$ endlich?

O.E. können wir annehmen, daß G in Chomsky-Normalform ist.

Wir definieren einen Graphen (W, E) auf der Menge der produktiven Nichtterminale mit folgender Kantenrelation:

$$E = \big\{(A, B) \in W \times W \;\big|\; \exists C \in W : (A \to BC) \in P \text{ oder } (A \to CB) \in P\big\}$$

Beobachtung: $(A, B) \in E$ gilt genau dann, wenn es einen vollständigen A-Ableitungsbaum gibt, so daß B ein Kind der Wurzel beschriftet.

Behauptung: $|L(G)| = \infty \iff \exists A \in W\colon (S, A) \in E^*$ und $(A, A) \in E^+$.

Beachte: $(B, C) \in E^*$ (bzw. $(B, C) \in E^+$) bedeutet hierbei, daß es einen Pfad (bzw. einen nicht leeren Pfad) von B nach C in der Relation E gibt.

Beweis:
„$\Leftarrow$“: Sei $A \in W$ mit $(S, A) \in E^*$ und $(A, A) \in E^+$.

$\rightsquigarrow$ Für jedes $n \in \mathbb{N}$ existiert ein vollständiger Ableitungsbaum T_n mit einem Ast, auf dem A wenigstens n mal vorkommt.

$\rightsquigarrow$ Es gibt unendlich viele vollständige Ableitungsbäume.

Da G in Chomsky-Normalform ist, kann jedes Wort $w \in \Sigma^*$ nur endlich viele Ableitungsbäume haben (sie haben alle die Größe $3|w|$).

$\rightsquigarrow |L(G)| = \infty$

„$\Rightarrow$“: Sei $L(G)$ unendlich.

Angenommen, es gäbe kein $A \in W$ mit $(S, A) \in E^*$ und $(A, A) \in E^+$.

$\rightsquigarrow$ jeder in S startende Pfad in (W, E) enthält $\leq |W|$ Kanten

$\rightsquigarrow$ jeder Ableitungsbaum von G hat nur Äste der Länge $\leq |W|$

$\rightsquigarrow$ es gibt nur endlich viele Ableitungsbäume, d.h. $|L(G)| < \infty$, im Widerspruch zur Annahme. $\square$

Beispiel:

G ist die Grammatik in Chomsky-Normalform mit den folgenden Produktionen:

$$\begin{aligned} S &\rightarrow AC \\ A &\rightarrow BC \\ B &\rightarrow CA \mid b \\ C &\rightarrow a \end{aligned}$$

Dann gilt $W = \{A, B, C, S\}$ (vgl. Folie 16.4) und die Kantenrelation E ist

$$E = \big\{(S, A), (S, C), (A, B), (A, C), (B, C), (B, A)\big\}.$$

Damit gilt $S \xrightarrow{E} A \xrightarrow{E} B \xrightarrow{E} A$.

Also ist $L(G)$ unendlich.

Unentscheidbarkeit bei kontextfreien Sprachen

Folgende Probleme sind für kontextfreie Sprachen nicht entscheidbar, d.h. es gibt kein entsprechendes Verfahren (vgl. Abschnitt „Unentscheidbare Probleme“):

- Universalitätsproblem: Gegeben eine kontextfreie Grammatik G. Gilt $L(G) = \Sigma^*$?
- Äquivalenzproblem: Gegeben zwei kontextfreie Grammatiken G_1 und G_2. Gilt $L(G_1) = L(G_2)$?
- (Inhärente) Mehrdeutigkeit: Gegeben eine kontextfreie Grammatik G. Ist G (inhärent) mehrdeutig?
- Komplementierbarkeit: Gegeben eine kontextfreie Grammatik G. Ist $\Sigma^* \setminus L(G)$ kontextfrei?
- Determinisierbarkeit: Gegeben ein PDA M. Existiert ein DPDA M' mit $L(M) = L(M')$?
- Regularität: Gegeben ein PDA M. Ist $L(M)$ regulär?

Entscheidbarkeit bei deterministisch kontextfreien Sprachen

Folgende Probleme sind entscheidbar:

- Regularität: Gegeben ein deterministischer Kellerautomat P. Ist $L(P)$ regulär?
- Äquivalenzproblem: Gegeben zwei deterministische Kellerautomaten P_1 und P_2. Gilt $L(P_1) = L(P_2)$?
- Universalitätsproblem: Gegeben ein deterministischer Kellerautomat P. Gilt $L(P) = \Sigma^*$?

Bemerkung: Die anderen Probleme der vorherigen Folie sind für DPDAs trivial:

- Jede deterministisch kontextfreie Sprache hat eine eindeutige Grammatik.
- Das Komplement jeder deterministisch kontextfreien Sprache ist wieder kontextfrei.
- Zu jedem DPDA existiert ein äquivalenter DPDA.

Unentscheidbarkeit bei deterministisch kontextfreien Sprachen

Folgende Probleme sind für deterministische Kellerautomaten nicht entscheidbar:

- Inklusionsproblem: Gegeben zwei deterministische Kellerautomaten P_1 und P_2. Gilt $L(P_1) \subseteq L(P_2)$?
- Schnittproblem: Gegeben zwei deterministische Kellerautomaten P_1 und P_2. Gilt $L(P_1) \cap L(P_2) = \emptyset$?

Damit sind diese Probleme auch für allgemeine kontextfreie Sprachen unentscheidbar.

Das folgende Schnittproblem ist jedoch entscheidbar:

- Gegeben eine kontextfreie Grammatik G_1 und eine rechtslineare Grammatik G_2. Gilt $L(G_1) \cap L(G_2) = \emptyset$?

Entscheidungsverfahren:

1. Dann können ein PDAE M_1 und ein NFA A_2 berechnet werden mit $L(M_1) = L(G_1)$ und $L(A_2) = L(G_2)$.
2. Hieraus kann ein PDAE M konstruiert werden mit $L(M) = L(M_1) \cap L(A_2) = L(G_1) \cap L(G_2)$.
3. Der PDAE M kann dann in einen PDA M' umgewandelt werden mit $L(M) = L(M')$.
4. Der PDA M' kann dann in eine kontextfreie Grammatik G umgewandelt werden mit $L(M') = L(G)$.
5. Durch Bestimmung der produktiven Nichtterminale von G kann dann ermittelt werden, ob S nicht-produktiv ist und damit, ob $L(G) = L(G_1) \cap L(G_2)$ leer ist.

Zusammenfassung kontextfreie Sprachen

- durch kontextfreie Grammatiken erzeugt, durch PDAs bzw. PDAEs erkannt, LL-Parsing
- Chomsky- und Greibach-Normalform, Pumping-Lemma, CYK-Algorithmus
- Klasse der $\sim$ abgeschlossen unter Vereinigung, Verkettung, Iteration, Schnitt mit reg. Sprachen;
  nicht abgeschlossen unter Komplement und Schnitt
- Algorithmen für Wortproblem, Leerheit, Endlichkeit;
  es gibt keinen Algorithmus für Äquivalenz usw.
- Klasse der det. kontextfreien Sprachen abgeschlossen unter Komplement,
  es gibt Algorithmus für Äquivalenz,
  es gibt keine Algorithmen für Schnitt- und Inklusionsproblem