Übung 6: CNF und CFL-Pumping Lemma

Theoretische Informatik Sommersemester 2013

Markus Kaiser

July 11, 2013

Definition (Chomsky-Normalform)

Eine kontextfreie Grammatik ist in Chomsky-Normalform (CNF) genau dann wenn alle Produktionen die Form

$$A \to a \quad \text{oder} \quad A \to BC$$

haben.

Satz

Zu jeder CFG G existiert eine CFG G′ in Chomsky-Normalform mit

$$L(G') = L(G) \setminus \{\epsilon\}$$

CNF Konstruktion

Idee

Sei $G = (V, \Sigma, P, S)$ eine CFG.

1. Eliminiere ϵ-Produktionen
2. Eliminiere Kettenproduktionen
3. Ersetze Terminale durch Nichtterminale
4. Verkürze Ketten von Nichtterminalen der Länge ≥ 3

Idee

Sei $G = (V, \Sigma, P, S)$ eine CFG.

1. Eliminiere ϵ-Produktionen
2. Eliminiere Kettenproduktionen
3. Ersetze Terminale durch Nichtterminale
4. Verkürze Ketten von Nichtterminalen der Länge ≥ 3

Sind $B \to \epsilon$ und $A \to \alpha B \beta$ in P, dann füge $A \to \alpha\beta$ hinzu.
Entferne danach alle ϵ-Produktionen.

$$S \to Ab, \quad A \to aAA \mid \epsilon$$

neu:

$$S \to b$$
$$A \to aA \mid a$$

Idee

Sei $G = (V, \Sigma, P, S)$ eine CFG.

1. Eliminiere ϵ-Produktionen
2. Eliminiere Kettenproduktionen
3. Ersetze Terminale durch Nichtterminale
4. Verkürze Ketten von Nichtterminalen der Länge ≥ 3

Sind $A \rightarrow B$ und $B \rightarrow \alpha$ in P, dann füge $A \rightarrow \alpha$ hinzu. Entferne danach alle Kettenproduktionen und unerreichbaren Symbole.

$$S \rightarrow A, \quad A \rightarrow a \mid B, \quad B \rightarrow bS$$

neu:

$$A \rightarrow a \mid bS$$
$$S \rightarrow a \mid bS$$

CNF Konstruktion

Idee

Sei $G = (V, \Sigma, P, S)$ eine CFG.

1. Eliminiere ϵ-Produktionen
2. Eliminiere Kettenproduktionen
3. Ersetze Terminale durch Nichtterminale
4. Verkürze Ketten von Nichtterminalen der Länge ≥ 3

Ersetze jedes $a \in \Sigma$ in einer rechten Seite länger als 1 durch ein neues Nichtterminal.

$$S \to aa \mid Bb \mid b, \quad B \to \dots$$

neu:

$$S \to X_a X_a \mid B X_b \mid b$$
$$X_a \to a, \quad X_b \to b$$

CNF Konstruktion

Idee

Sei $G = (V, \Sigma, P, S)$ eine CFG.

1. Eliminiere ϵ-Produktionen
2. Eliminiere Kettenproduktionen
3. Ersetze Terminale durch Nichtterminale
4. Verkürze Ketten von Nichtterminalen der Länge ≥ 3

Ersetze jede Produktion der Form $A \to B_1 B_2 \dots B_k$ durch neue Nichtterminale mit Produktionen der Länge 2.

$$S \to X_a X_b B X_a, \quad X_a \to a, \quad X_b \to b, \quad B \to \dots$$

neu:

$$S \to X_a T_1$$
$$T_1 \to X_b T_2, \quad T_2 \to B X_a$$

Eigenschaften von Symbolen

Definition

Sei $G = (V, \Sigma, P, S)$ eine CFG.
Ein Symbol $X \in V \cup \Sigma$ ist

nützlich	es gibt $S \rightarrow^*_G w \in \Sigma^*$ in der X vorkommt
erzeugend	es gibt $X \rightarrow^*_G w \in \Sigma^*$
erreichbar	es gibt $S \rightarrow^*_G \alpha X \beta$

Satz

Nützliche Symbole sind erzeugend und erreichbar. Aber nicht notwendigerweise umgekehrt.

$$S \rightarrow AB \mid a, \quad A \rightarrow b$$

Satz (Pumping Lemma für kontextfreie Sprachen)

Sei $L \subseteq \Sigma^*$ kontextfrei. Dann gibt es ein $n > 0$, so dass sich jedes $z \in L$ mit $|z| \geq n$ so in $z = uvwxy$ zerlegen lässt, dass

- $vx \neq \epsilon$
- $|vwx| \leq n$
- $\forall i \geq 0. uv^iwx^iy \in L$