Dr. Sebastian Bab WiSe 12/13

Theoretische Grundlagen der Informatik für TI

Termin: VL 13 vom 29.11.2012

Kontextfreie Grammatiken, Chomsky-Normalform, Pumping-Lemma für kontextfreie Sprachen

- $L = \{ a^n b^n c^m \mid n, m \in \mathbb{N}^+ \}$ ist nicht regulär

  Beweis:
  Beweis mit Pumpinglemma: Wäre L regulär gäbe es eine Pumpingzahl n. Das Wort $w = a^n b^n m^1 = a^n b^n m \in L$ wäre dann Wort der Sprache mit $|w| = 2n + 1 \geq n$. Man kann also laut Pumpinglemma $w = xyz$ so zerlegen, dass $x = a^p$, $y = a^q$ und $z = a^{n-p-q} b^n c$, $1 \leq q \leq n$, $0 \leq p \leq n$ und $p + q \leq n$ gilt und dann $xy^0 z = xz \in L$ gilt. $q \geq 1$ gilt, hat xz mitdestens ein a weniger als xyz, während die Anzahl der b unverändert ist. Also erfüllt xz nicht die Eigenschaften der Sprache L und L kann nicht regulär sein. $\square$

- $L = \{ a^n b^n c^m \mid n, m \in \mathbb{N}^+ \}$ ist kontextfrei

  Beweis:
  Beweis durch Angabe einer kontextfreien Grammatik G mit $L(G) = L$:

$$\begin{array}{lcl} S & \to & AB \\ A & \to & ab|aAb \\ B & \to & c|Bc \end{array}$$

$\square$

- kontextfreie Sprachen sind nicht so leicht durch Algorithmen zu bearbeiten, wie reguläre Sprachen (DEA's reichen z.B. nicht)
- Gibt es eine algorithmenfreundliche Form für kontextfreie Grammatiken?

Definition 1 (Chomski-Normalform)
Eine kontextfreie Grammatik $G = (V, \Sigma, S, P)$ ist in Chomski-Normalform (CNF)*, wenn alle Regeln $\alpha \to \beta$ die folgende Form haben:*

- $\alpha \in V$
- $\beta \in \Sigma \cup (V \circ V)$

- Beispiel: G_2 ist in CNF

$$\begin{array}{rcl} S & \to & AB \\ A & \to & AB|a \\ B & \to & BC|b \\ C & \to & a|b \end{array}$$

- Grammatiken in CNF sind wesentlich leichter algorithmisch zu bearbeiten, da sie nützliche Eigenschaften haben.

Satz 2
Zu jeder kontextfreien Grammatik G gibt es eine Grammatik G' in CNF, die die selbe Sprache erzeugt.

- Für $L_1 = \{\, a^n b^n c^m \mid m, n \in \mathbb{N}^+ \,\}$ und $L_2 = \{\, a^m b^m c^m \mid m \in \mathbb{N}^+ \,\}$ wissen wir, dass sie nicht regulär sind. Für die Sprache L_1 wissen wir außerdem schon, dass sie kontextfrei ist. Ist L_2 kontextfrei? Wie lässt sich das beweisen oder widerlegen?
- Um zu zeigen, dass Sprachen nicht kontextfrei sind, gibt es auch ein Pumpinglemma für kontextfreie Sprachen.

Satz 3
(Pumpinglemma) Sei L eine kontextfreie Sprache. Dann existiert eine Konstante n (abhängig von L), so dass für alle Worte $z \in L$ mit $|w| \geq n$ gilt, dass es eine Zerlegung $z = uvwxy$ gibt, die folgende Eigenschaften hat:

1. $|vx| > 0$
2. $|vwx| \leq n$
3. *Für alle $k \in \mathbb{N}$ gilt: $uv^k wx^k y \in L$.*

- $L = \{\, a^m b^m c^m \mid m \in \mathbb{N}^+ \,\}$ ist nicht kontextfrei.

  Beweis:
  Angenommen, $L = \{\, a^m b^m c^m \mid m \in \mathbb{N}^+ \,\}$ wäre kontextfrei. Dann gibt es also eine Zahl n, ab der für alle Wörter die Pumpingeigenschaften gelten. Sei also $a^n b^n c^n \in L$

ein Wort der Sprache. Da $|a^n b^n c^n| = 3n \geq n$ müssen für dieses Wort die Pumpingeigenschaften gelten. Also gibt es eine Zerlegung

$$a^n b^n c^n = uvwxy$$

mit $|vwx| \leq n$ und $|vx| > 0$. Also kommt in vwx mindestens eines der Zeichen a, b, c nicht vor, da dazu ein längeres Wort nötig wäre. Laut Pumpinglemma muss nun auch das Wort $uv^0wx^0y = uwy \in L$ sein. Also wurde mindestens ein a, b oder c entfernt, und für mindestens eines der Zeichen a, b oder c blieb die Anzahl der Zeichen unverändert. Damit kann es nicht mehr gleich viele a, b und c in uwy geben und das Wort erfüllt nicht die Eigenschaften der Sprache L. Also gilt das Pumpinglemma für L nicht und daher kann L nicht kontextfrei sein. $\square$