16 Pumping-Lemma für kontextfreie Sprachen

In diesem Kapitel wird ein Pumping-Lemma für kontextfreie Sprachen formuliert, auf dessen Beweis hier allerdings verzichtet wird. Analog zum Pumping-Lemma für reguläre Sprachen in Kapitel 7[4] eignet es sich zum Nachweis, dass bestimmte Sprachen nicht kontextfrei sind.

Theorem 11
Zu jeder kontextfreien Sprache L existiert eine natürliche Zahl $p \in \mathbb{N}$, so dass gilt:

Ist $z \in L$ mit $length(z) \geq p$, dann lässt sich z schreiben als $z = uvwxy$, wobei $length(vwx) \leq p$ ist und $vx \neq \lambda$, und für alle $i \in \mathbb{N}$ gilt:

$$uv^i wx^i y \in L.$$

Mit dem Pumping Lemma für kontextfreie Sprachen lässt sich z.B. zeigen, dass die Sprache $L_0 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ nicht kontextfrei ist. Dieses Ergebnis kann wiederum benutzt werden, um zu zeigen, dass die Klasse der kontextfreien Sprachen nicht abgeschlossen gegenüber Durchschnitt ist.

Korollar 12

1. *Die Sprache $L_0 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ ist nicht kontextfrei.*
2. *Die Klasse der kontextfreien Sprachen ist nicht abgeschlossen gegenüber Durchschnitt, d.h. es gibt kontextfreie Sprachen, deren Durchschnitt nicht kontextfrei ist.*

Beweis.

1. Angenommen, L_0 wäre kontextfrei. Sei p die zu L_0 gehörende Zahl aus dem Pumping-Lemma. Dann gibt es für $n \in \mathbb{N}$ mit $3n \geq p$ eine Zerlegung $a^n b^n c^n = uvwxy$, derart dass $uv^i wx^i y \in L$ für alle $i \leq p$, $length(vwx) \leq p$ und $vx \neq \lambda$.
   Enthielte v oder x a's und b's oder b's und c's, so käme in uv^2wx^2y im Widerspruch zur Definition von L_0 ein a hinter einem b bzw. ein b hinter einem c vor. Also ist $v = a^r$ oder $v = b^r$ oder $v = c^r$ und entsprechend $x = c^s$ oder $x = b^s$ oder $x = a^s$, wobei $r + s \geq 1$. Da v in z stets vor x liegt, sind das insgesamt sechs Fälle. Im Fall $v = a^r$ und $x = c^s$ gilt $u = a^j$, $w = a^k b^n c^l$, $y = c^m$ und $z = uvwxy = a^j a^r a^k b^n c^l c^s c^m$ mit $j + r + k = n = l + s + m$. Damit erhält man $uv^0wx^0y = a^j \lambda a^k b^n c^l \lambda c^m = a^{n-r} b^n c^{n-s} \in L_0$, was wegen $r + s \geq 1$ der Definition von L_0 widerspricht. Analog führen die anderen fünf Fälle zum Widerspruch.
2. Die durch die Produktionen $S_1 ::= S_1 c \mid A$ und $A ::= aAb \mid \lambda$ sowie $S_2 ::= aS_2 \mid B$ und $B ::= bBc \mid \lambda$ definierten Grammatiken erzeugen die kontextfreien Sprachen $L_1 = \{a^m b^m c^n \mid m, n \in \mathbb{N}\}$ und $L_2 = \{a^m b^n c^n \mid m, n \in \mathbb{N}\}$, deren Durchschnitt gerade die nicht kontextfreie Sprache L_0 aus Punkt 1 ist. □

[4] Dort ist das Pumping-Lemma für die von endlichen Automaten erkannten Sprachen formuliert. Reguläre Sprachen sind genau die von endlichen Automaten erkannten; Regularität ist aber die viel griffigere Bezeichnung und wird deshalb hier verwendet.