SS 2015

Zentralübung zur Vorlesung Theoretische Informatik

Dr. Werner Meixner

Fakultät für Informatik
TU München

http://www14.in.tum.de/lehre/2015SS/theo/uebung/

7. Mai 2015

ZÜ IV

Übersicht:

1. Übungsbetrieb Fragen, Probleme?
2. Thema Pump Lemma
3. Vorbereitung TA Blatt 4

1. Fragen, Anregungen?

Aktuelle Fragen?

Achtung Terminänderung:
Blatt 7 wird am 28. Mai ausgegeben und
am 28. Mai findet eine Zentralübung statt.

2. Thema: Pump Lemma

Das Pumping Lemma wird in der VL in drei Varianten behandelt:

1. Pumping Lemma für reguläre Sprachen
2. Pumping Lemma für kontextfreie Sprachen
3. Ogden-Lemma für kontextfreie Sprachen

In jedem Fall ist das Lemma eine Aussage über die Abgeschlossenheit der Sprachen gegenüber gewissen unendlichen Mengen mit einer gewissen iterativen Struktur.

Allerdings gibt es nichtreguläre bzw. nichtkontextfreie Sprachen, die diese Abgeschlossenheitseigenschaften ebenfalls besitzen.

Sei $L \subseteq \Sigma^*$ regulär bzw. kontextfrei. Dann gibt es eine reguläre bzw. kontextfreie Pumping-Lemma-Zahl $n \in \mathbb{N}$,

d.h., dass für alle $z \in L$ mit $|z| \geq n$ gilt:

z ist zerlegbar in $z = wxy$, so dass:	bzw.	z ist zerlegbar in $z = uvwxy$, so dass:
1. $|x| \geq 1$		1. $|vx| \geq 1$
2. $|wx| \leq n$		2. $|vwx| \leq n$
und		und
3. $\forall i \geq 0 : \quad wx^iy \in L$		3. $\forall i \geq 0 : \quad uv^iwx^iy \in L$

Wir definieren eine Kennzeichnung M von Vorkommen von Buchstaben in einem Wort w als

Markierung von m Vorkommen.

Eine Markierung von 3 Buchstaben in $w = abbaccd$ ist beipielsweise durch $w = abbaccd$ gegeben.

Die Länge $|w|_M$ eines Wortes w bezüglich einer Markierung M ist dann die Anzahl von Markierungen in w.

Beispiel: $|abbaccd|_M = 3$.

Ogden-Lemma:

Sei $L \subseteq \Sigma^*$ kontextfrei. Dann
gibt es eine Zahl $n \in \mathbb{N}$, so dass

für alle $z \in L$ mit $|z| \geq n$ und
für alle Markierungen M von $m \geq n$ Buchstaben in z:

z ist zerlegbar in $z = uvwxy$, so dass:	z.Vgl.	z ist zerlegbar in $z = uvwxy$, so dass:
1. $\|vx\|_M \geq 1$		1. $\|vx\| \geq 1$
2. $\|vwx\|_M \leq n$		2. $\|vwx\| \leq n$
und		und
3. $\forall i \geq 0 : \quad uv^iwx^iy \in L$		3. $\forall i \geq 0 : \quad uv^iwx^iy \in L$

3. Vorbereitung TA Blatt 4

3.1 VA 1

Gegeben sei die Sprache $L = \{a^n \mid n = 2^k, k \in \mathbb{N}\}$. Zeigen Sie, dass L nicht regulär ist.

Pump Lemma für reguläre Sprachen:

Sei $L \subseteq \Sigma^*$ regulär. Dann gibt es eine reguläre
Pump Lemma-Zahl $n \in \mathbb{N}$,
d.h., dass für alle $z \in L$ mit $|z| \geq n$ gilt:

z ist zerlegbar in
$z = wxy$,
so dass:

1. $|x| \geq 1$
2. $|wx| \leq n$

und

3. $\forall i \geq 0 : \quad wx^iy \in L$

Lösung

Wir beweisen durch Widerspruch und nehmen an, dass L regulär ist.

Sei $n \geq 1$ eine Pumping-Lemma-Zahl für L. Wir nehmen sofort auch $n > 2$ an, was einerseits zulässig ist und andererseits die unten benötigte Ungleichung $2n < 2^n$ gültig macht.

Wir betrachten $z = a^{2^n}$.
Dann gilt $|z| \geq n$ und $z \in L$.

Nach Pumping-Lemma können wir z wie folgt zerlegen.

Sei $z = uvw$ mit $|uv| \leq n$ und $v \neq \epsilon$,
so dass für alle $i \in \mathbb{N}_0$ gilt $uv^iw \in L$.

Zunächst gilt $v = a^r$ mit $1 \leq r \leq n$. Es folgt $uw = a^{2^n - r}$ und $uw \in L$, d.h. $a^{2^n-r} = a^{2^k}$ und mithin $2^n - r = 2^k$ für ein existierendes $k \in \mathbb{N}$.

Dies ist ein Widerspruch zur Tatsache, dass aus $r \leq n < 2^{n-1}$ die Ungleichungen $0 < 2^{n-1} - r$ und mithin
$2^{n-1} < 2 \cdot 2^{n-1} - r = 2^n - r < 2^n$ folgen.

3.2 VA 2

Für formale Sprachen L gibt es eine Äquivalenz $\equiv_L$ über L. Diese Äquivalenzrelation hängt bei regulären Sprachen eng mit dem Begriff der Äquivalenz von Zuständen deterministischer endlicher Automaten zusammen.

Sei $M = (Q, \Sigma, \delta, s_0, F)$ eine DFA.

Definition (Äquivalenz von Zuständen $p, q \in Q$)

$$p \equiv_M q \quad \Longleftrightarrow \quad \forall w \in \Sigma^* : (\hat{\delta}(p, w) \in F \Leftrightarrow \hat{\delta}(q, w) \in F)\,.$$

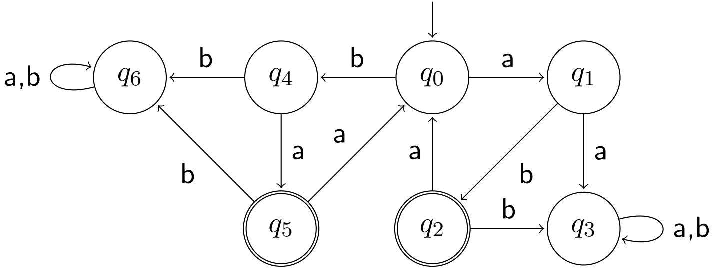
Es gilt

$$p \equiv_M q \quad :\Longleftrightarrow \quad L_M(p) = L_M(q)\,.$$

mit der Definition

$$L_M(p) \quad := \quad \{w \in \Sigma^*\,;\, \hat{\delta}(p, w) \in F\}\,.$$

Sei $\Sigma = \{a, b\}$. Der Automat M sei durch das folgende Diagramm gegeben. Zeigen Sie $q_3 \not\equiv_M q_4$ und $q_3 \equiv_M q_6$.



Lösung

$\hat{\delta}(q_3, a) = q_3 \notin F$, aber $\hat{\delta}(q_4, a) = q_5 \in F$.

Damit ist die Definition von Äquivalenz nicht erfüllt,
d.h. es gilt $q_3 \not\equiv_M q_4$.

Alle von q_3 bzw. q_6 ausgehenden Übergänge führen wieder nach q_3 bzw. q_6.

Das heißt, für alle $w \in \Sigma^*$ gilt
$\hat{\delta}(q_3, w) = q_3 \notin F$ und $\hat{\delta}(q_6, w) = q_6 \notin F$

und damit ist $q_3 \equiv_M q_6$.

3.3 VA 3

Wir nennen eine Phrasenstrukturgrammatik $G = (V, \Sigma, P, S)$ *nullierbar kontextfrei*, wenn alle Regeln aus P die folgende Form besitzen.

$$A \underset{P}{\rightarrow} \alpha \qquad \text{mit} \quad A \in V,\ \Gamma = V \cup \Sigma,\ \alpha \in \Gamma^* .$$

Γ^* heißt Menge der Satzformen über dem Vokabular Γ.

Sei G eine nullierbar kontextfreie Grammatik.

1. Man zeige für alle $u, v, w \in \Gamma^*$ die Zerlegungseigenschaft

$$uv \underset{G}{\rightarrow} w \quad \Longrightarrow$$
$$\exists\, u', v' \in \Gamma^* : \quad u \underset{G}{\rightarrow}^* u' \,\wedge\, v \underset{G}{\rightarrow}^* v' \,\wedge\, u'v' = w\,.$$

2. Es gilt für alle $u, v \in \Gamma^*$, $a \in \Sigma$ und $n \in \mathbb{N}_0$

$$uv \underset{G}{\rightarrow}^* a^n \quad \Longrightarrow$$
$$\exists\, p, q \in \mathbb{N}_0 : \quad p + q = n \wedge u \underset{G}{\rightarrow}^* a^p \wedge v \underset{G}{\rightarrow}^* a^q\,.$$

1. Man zeige für alle $u, v, w \in \Gamma^*$ die Zerlegungseigenschaft

$$uv \underset{G}{\rightarrow} w \quad \Longrightarrow$$
$$\exists\, u', v' \in \Gamma^* : \quad u \underset{G}{\rightarrow}^* u' \ \wedge\ v \underset{G}{\rightarrow}^* v' \ \wedge\ u'v' = w\,.$$

Lösung

Wenn eine Produktion $A \underset{P}{\rightarrow} \alpha$ zur Anwendung auf uv kommt, dann ist die Anwendungsstelle entweder in u oder in v (weil knf).

Falls die Anwendungsstelle in u liegt, dann gibt es $x, y \in \Gamma^*$, so dass gilt $u = xAy$. Dann folgt $w = x\alpha yv$.
Daraus folgt aber mit $u' = x\alpha y$ und $v' = v$, dass gilt $u \underset{G}{\rightarrow} u'$ und $v \underset{G}{\rightarrow}^* v'$, und insbesondere auch $u'v' = w$.

Entsprechendes gilt, falls die Anwendungsstelle der Produktion in v liegt.

Damit ist die Zerlegungseigenschaft bewiesen.

2. Es gilt für alle $u, v \in \Gamma^*$, $a \in \Sigma$ und $n \in \mathbb{N}_0$

$$uv \xrightarrow[G]{}^* a^n \quad \Longrightarrow$$
$$\exists\, p, q \in \mathbb{N}_0 : \quad p + q = n \land u \xrightarrow[G]{}^* a^p \land v \xrightarrow[G]{}^* a^q\,.$$

Lösung

Die Aussage folgt leicht durch Induktion über die Länge k der Ableitung für $uv \xrightarrow[G]{}^* a^n$.

Für $k = 0$ gilt $uv = a^n$. Daraus folgt aber unmittelbar $u = a^p$ und $v = a^q$ für geeignete $p, q \in \mathbb{N}_0$ mit $p + q = n$. Wegen Reflexivität gilt dann trivialerweise $u \xrightarrow[G]{}^* a^p$ und $v \xrightarrow[G]{}^* a^q$.

Von k auf $k+1$ schließen wir für alle $k \geq 0$ wie folgt:

Falls $uv \underset{G}{\rightarrow}^* a^n$ durch eine Ableitung der Länge $k+1$ dargestellt wird, dann gilt $uv \underset{G}{\rightarrow} w \underset{G}{\rightarrow}^* a^n$, wobei $w \underset{G}{\rightarrow}^* a^n$ durch eine Ableitung der Länge k dargestellt wird.

Wegen der in Teilaufgabe 1 bewiesenen Zerlegungseigenschaft können wir für gewisse u', v' schreiben

$$uv \underset{G}{\rightarrow} u'v' = w \underset{G}{\rightarrow}^* a^n \quad \text{mit} \quad u \underset{G}{\rightarrow}^* u' \quad \text{und} \quad v \underset{G}{\rightarrow}^* v'.$$

Nach Induktionsannahme folgt $u' \underset{G}{\rightarrow}^* a^p$ und $v' \underset{G}{\rightarrow}^* a^q$, mithin $u \underset{G}{\rightarrow}^* a^p$ und $v \underset{G}{\rightarrow}^* a^q$ für gewisse $p, q \in \mathbb{N}_0$.

3.4 VA 4

Welche Symbole einer durch die folgenden Produktionen gegebenen Grammatik sind erzeugend, welche erreichbar, und welche nützlich?

$$
\begin{array}{lllllll}
S & \rightarrow & A \mid a\,, & & & \\
A & \rightarrow & BC\,, & B & \rightarrow & AE \mid C\,, \\
C & \rightarrow & aS\,, & D & \rightarrow & aB \mid FS \\
E & \rightarrow & EA\,, & F & \rightarrow & Ca \mid a\,.
\end{array}
$$

Definitionen

Ein Nichtterminal A heißt

1. *erreichbar* gdw es eine Ableitung $S \underset{G}{\rightarrow}^* \alpha A \beta$ gibt,
2. *erzeugend* gdw es eine Ableitung $A \underset{G}{\rightarrow}^* w \in \Sigma^*$ gibt,
3. *nützlich* gdw es eine Ableitung $S \underset{G}{\rightarrow}^* \alpha A \beta \underset{G}{\rightarrow}^* w \in \Sigma^*$ gibt.

Nützliche Nichtterminale sind erreichbar und erzeugend.
Die Umkehrung gilt nicht notwendigerweise (Beweis?)

Lösung:

$$\begin{array}{rclcrcl} S & \rightarrow & A \mid a\,, & & & & \\ A & \rightarrow & BC\,, & & B & \rightarrow & AE \mid C\,, \\ C & \rightarrow & aS\,, & & D & \rightarrow & aB \mid FS \\ E & \rightarrow & EA\,, & & F & \rightarrow & Ca \mid a\,. \end{array}$$

Alle Symbole außer E sind erzeugend.

Alle Symbole außer D und F sind erreichbar.

Die anderen Symbole sind nützlich, wie die folgende Ableitung zeigt:

$$S \rightarrow A \rightarrow BC \rightarrow CC \rightarrow^2 aSaS \rightarrow^2 aaaa$$