

Pumping Lemma: 3. Anwendung (1)

Schließlich wollen wir noch die Sprache der *unären Primzahlen* untersuchen:

$$L = \{a^n \mid n \text{ ist eine Primzahl}\}$$

Auch diese Sprache ist nicht regulär, und wieder führen wir den Beweis mit dem Pumping Lemma.

Es sei $n \in \mathbb{N}$ gegeben. Wir wählen eine Primzahl $p > n + 2$. Eine solche existiert, denn bekanntlich gibt es unendlich viele Primzahlen.

Damit ist $a^p \in L$ und $|a^p| = p > n$. Nun müssen wir wieder für eine beliebige Zerlegung $a^p = uvw$ mit $|v| \geq 1$ und $|uv| \leq n$ nachweisen, dass es ein $i \geq 0$ gibt, für das $uv^i w \notin L$ gilt.

3. Anwendung (2)

Es sei k die Länge von v , also $k = |v|$.

Nun setzen wir $i = p + 1$ und untersuchen $uv^i w$, also das Wort $uv^{p+1}w$, dessen Länge offenbar so zu berechnen ist:

$$\begin{aligned} |uv^{p+1}w| &= \\ |u| + (p+1)|v| + |w| &= p \cdot |v| + |u| + |v| + |w| = \\ p \cdot k + |uvw| &= p \cdot k + p = p \cdot (k+1) \end{aligned}$$

Wir stellen fest, dass $p > 1$ (sogar $p > n + 2$) und $k + 1 > 1$, da $k \geq 1$ nach Voraussetzung. Also ist die Länge von $uv^{p+1}w$ keine Primzahl, d.h. dieses Wort gehört nicht zu L .

Übrigens ist auch diese Sprache keine Typ-2 Sprache. Der Grund ist im wesentlichen der, dass es überhaupt keine *unäre* Typ-2 Sprache gibt, die nicht schon Typ-3 ist. Wir werden das im Abschnitt über die kontextfreien Sprachen beweisen.

Die Myhill-Nerode Äquivalenz

Wir betrachten die folgende Äquivalenzrelation R_L :

$$x R_L y \iff [\forall w \in \Sigma^* : xw \in L \iff yw \in L]$$

Die notwendigen Bedingungen für eine Äquivalenzrelation (Reflexivität, Symmetrie und Transitivität) sind leicht zu zeigen.

Wenn L Typ-3 Sprache ist, also $L = T(M)$ für einen DEA M , dann ist auch die folgende Relation R_M eine Äquivalenzrelation:

$$x R_M y \iff \hat{\delta}(z_0, x) = \hat{\delta}(z_0, y)$$

Es gilt:

$$x R_M y \implies \forall w : \hat{\delta}(z_0, xw) = \hat{\delta}(z_0, yw) \implies x R_L y$$

also ist R_M eine *Verfeinerung* von R_L (d.h. jede Äquivalenzklasse bzgl. R_M ist komplett in einer Äquivalenzklasse bzgl. R_L enthalten).

Der Satz von Myhill und Nerode

Als *Index* einer Äquivalenzrelation bezeichnet man die Zahl der durch die Relation induzierten Äquivalenzklassen. Der Index kann eine natürliche Zahl sein, er kann aber auch unendlich sein.

Satz: Die Sprache $L \subseteq \Sigma^*$ ist genau dann regulär, wenn der Index der Äquivalenzrelation R_L endlich ist.

Was müssen wir für den Beweis tun?

Zunächst werden wir zeigen, dass für jede Typ-3 Sprache L der Index R_L endlich ist.

Danach zeigen wir, dass es für jede Sprache L mit endlichem Index R_L einen DEA M mit $T(M) = L$ gibt.

Beweis, 1. Richtung

Es sei also zunächst L eine Typ-3 Sprache, d.h. es existiert ein DEA M mit $L = T(M)$. Für die zu diesem DEA gehörende Äquivalenzrelation R_M gilt:

Wenn x und y in verschiedenen Äquivalenzklassen liegen, dann ist $\hat{\delta}(z_0, x) \neq \hat{\delta}(z_0, y)$.

Also kann der Index von R_M maximal die Anzahl der Zustände von M sein. Aber R_M ist eine Verfeinerung von R_L , d.h. der Index von R_L ist auch beschränkt durch die Anzahl der Zustände von M und daher endlich.

Damit ist die 1. Richtung bewiesen.

Beweis, 2. Richtung

Nun habe R_L endlichen Index. Wir repräsentieren die k Äquivalenzklassen durch je ein Wort x_i für $1 \leq i \leq k$. Dabei sei die Nummerierung so gewählt, dass $\varepsilon \in [x_1]$ gilt.

Setze $Z := \{[x_1], [x_2], \dots, [x_k]\}$.

Das heißt: Wir interpretieren die Äquivalenzklassen als Zustände des nachfolgend zu definierenden DEA $M = (Z, \Sigma, \delta, z_0, E)$:

- Z haben wir oben bereits angegeben.
- Σ ist das vorgegebene Alphabet.
- $\delta([x], a) = [xa]$ Ist δ dadurch wohldefiniert?
- $z_0 = [x_1]$
- $E = \{[x_i] \mid x_i \in L\}$ Ist E dadurch wohldefiniert?