Minimalautomaten

Frage: Ist der Äquivalenzautomat $A_{\equiv}$ der kleinste Automat ("*Minimalautomat*") der die Sprache $L(A)$ erkennt?

Minimalautomaten

Satz: Falls A keine unerreichbaren Zustände hat, ist $A_{\equiv}$ der kleinste Automat der die Sprache $L(A)$ erkennt

Idee: Übertrage die Relation $\equiv$ auf Sprachebene

Nerode-Relation

Def.: Sei $L \subset \Sigma^*$. Die Nerode Relation R_L ist wie folgt definiert:

$x,y \in \Sigma^* : xR_Ly \quad :\Leftrightarrow \quad (\forall z \in \Sigma^* : xz \in L \Leftrightarrow yz \in L)$

Bsp.: $L=0^*1^*$

- dann ist $\varepsilon \, R_L \, 0 \, R_L \, 00 \, R_L \, 000 \, R_L \, 000 \, R_L \dots$
  da $0^nz \in L \Leftrightarrow z \in L$
- weiter sind alle Wörter in Relation, in denen 1 vor 0 steht (keine Fortsetzung ist in L)
- weiter sind alle Wörter der Form 0^n1^m mit $m \geq 1$, $n \geq 0$ in Relation

→ R_L hat drei Äquivalenzklassen

$\Sigma^*/R_L = \{0^*, 0^*11^*, \Sigma^* \setminus \{0^* \cup 0^*11^*\}\} = \{[0]_{R_L}, [1]_{R_L}, [10]_{R_L}\}$

Eigenschaften der Nerode-Relation

Satz: R_L ist rechtsinvariante Äquivalenzrelation

- rechtsinvariant $xR_Ly \Rightarrow xzR_Lyz$
- reflexiv xR_Lx
- symmetrisch $xR_Ly \Rightarrow yR_Lx$
- transitiv $xR_Ly \wedge yR_Lz \Rightarrow xR_Lz$

Satz von Myhill-Nerode

Satz: L ist regulär $\Leftrightarrow$ R_L hat endlichen Index

Bem.: Index einer Äquivalenzrelation = #Äquivalenzklassen

Bsp.:

- $L = 0^*1^*$
  Index 3, da $\Sigma^*/R_L = \{[0]_{R_L}, [1]_{R_L}, [10]_{R_L}\}$
- $L = \{0^n1^n \mid n \geq 0\}$
  kein endlicher Index, da die Wörter 0^i1^j mit $i>j\geq 0$ paarweise nicht äquivalent sind
  (wir sehen später, dass L nicht regulär ist)

Beweis des Satzes von Myhill-Nerode

Bew.: R_L hat endlichen Index $\Rightarrow$ L ist dea Sprache

für dea $A=(Q,\Sigma,\delta,q_0,F)$ mit

$Q := \{[w]_{R_L} \mid w \in \Sigma^*\} = \Sigma^*/R_L$

$q_0 := [\varepsilon]_{R_L}$

$F := \{[w]_{R_L} \mid w \in L\}$

(wohldefiniert, da $wR_Lv \Rightarrow (w \in L \Leftrightarrow v \in L)$)

$\delta([w]_{R_L}, a) := [wa]_{R_L}$

(wohldefiniert, da $wR_Lv \Rightarrow waR_Lva$)

gilt nach Konstruktion $L(A)=L$

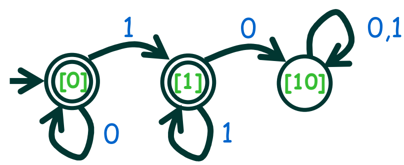
Bsp. zum Satz von Myhill-Nerode

$L=0^*1^*$

→ R_L hat drei Äquivalenzklassen

$\Sigma^*/R_L = \{[0]_{R_L}, [1]_{R_L}, [10]_{R_L}\}$

- $[0]_{R_L} = 0^*$
- $[1]_{R_L} = 0^*11^*$,
- $[10]_{R_L} = \Sigma^*\setminus\{0^*\cup 0^*11^*\}$



Beweis des Satzes von Myhill-Nerode

Bew.: L ist dea Sprache $\Rightarrow$ R_L hat endlichen Index

z.z. falls $L=L(A)$ mit dea $A=(Q,\Sigma,\delta,q_0,F)$, so ist $|\Sigma^*/R_L|\leq|Q|$

für $x,y\in\Sigma^*$ gilt

$$\delta^*(q_0,x) = \delta^*(q_0, y) \Rightarrow xR_Ly$$

oder

$$[x]_{R_L} \neq [y]_{R_L} \Rightarrow \delta^*(q_0,x) \neq \delta^*(q_0, y)$$

d.h.

$$|\Sigma^*/R_L| \leq |Q|$$

verschiedene Äquivalenzklassen liefern verschiedene Zustände

R_L vs. $\equiv$

Lemma: Falls $L=L(A)$ mit dea $A=(Q,\Sigma,\delta,q_0,F)$, so gilt für $A_\equiv=(Q_\equiv,\Sigma,\delta_\equiv,q_\equiv,F_\equiv)$, dass $|Q_\equiv| \leq |\Sigma^*/R_L|$

Bew.: für $x,y \in \Sigma^*$ gilt

$$xR_L y \Rightarrow \delta^*(q_0,x) \equiv \delta^*(q_0,y)$$

oder

$$\delta^*(q_0,x) \equiv^c \delta^*(q_0,y) \Rightarrow [x]_{R_L} \neq [y]_{R_L}$$

d.h.

$$|Q_\equiv| \leq |\Sigma^*/R_L|$$

verschiedene Äquivalenzklassen von $\equiv$ (Zustände von $A_\equiv$) liefern verschiedene Äquivalenzklassen von R_L

Minimalautomaten

Satz: Der Äquivalenzautomat $A_\equiv$ ist der kleinste Automat ("*Minimalautomat*") der die Sprache $L(A)$ erkennt

Bew.: sei A' Minimalautomat für $L(A)$ mit Zustandsmenge Q'

Myhill Nerode → $|\Sigma^*/R_L| \leq |Q'|$

Lemma → $|Q_\equiv| \leq |\Sigma^*/R_L|$

somit $|Q_\equiv| \leq |Q'|$, d.h. $A_\equiv$ ist minimal

Bem.: Minimalautomaten sind bis auf Isomorphie eindeutig !

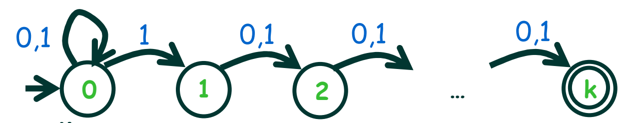
Anwendung des Satzes von Myhill-Nerode

Satz: $L_k = \{w\in\{0,1\}^* \mid k$ ter Buchstabe in w von hinten ist $1\}$ kann von einem nea mit $k+1$ Zuständen erkannt werden, aber jeder dea der L_k erkennt, hat mindestens 2^k Zustände

("exponentielles Wachstum bei Potenzmengenkonstruktion unvermeidbar")

Bew.:

- nea



- R_{L_k} hat 2^k Äquivalenzklassen :

  $u=u_1,\dots,u_k,\ v=v_1\dots v_k \in \{0,1\}^k$ mit $u_i \neq v_i$

  → nicht beide Fortsetzungen

  $u0^{i-1}$ und $v0^{i-1}$ sind gleichzeitig in L_k

Nachweis der Nichtregularität von Sprachen

Frage: Wie kann man zeigen, dass eine Sprache L nicht regulär ist?

Bsp.:

- $L_1 = \{0^n1^n \mid n\geq 0\}$

  regulär?

- $L_2 = \{w\in\{0,1\}^* \mid w$ hat die gleiche Zahl von Nullen und Einsen$\}$

  regulär?

Pumping-Lemma

Satz: Sei $L \subset \Sigma^*$ regulär.

Dann gibt es $n \geq 0$, so dass für alle $z \in L$ mit $|z| \geq n$ gilt:

es gibt $u,v,w \in \Sigma^*$, so dass $z=uvw$ mit $|uv| \leq n$ und $|v| \geq 1$, und

für alle $i>0$ ist $uv^iw \in L$

Beweis des Pumping-Lemma

Bew. ... : $L=L(A)$ mit dea $A=(Q,\Sigma,\delta,q_0,F)$ und $n := |Q|$

sei $z=z_1 \ldots z_m \in L$ mit $m=|z| \geq n$

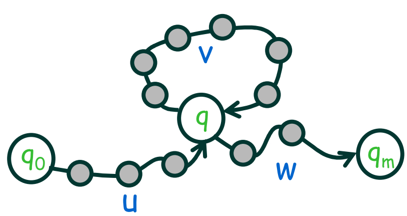
→ es gibt Folge von Zuständen $q_1,\ldots,q_m$ mit $q_m \in F$ und

$$q_{i+1} = \delta(q_i, z_{i+1}) \text{ für } 0 \leq i < m$$

→ $m+1$ Zustände werden durchlaufen,

→ manche Zustände werden mehrfach durchlaufen

("Schubfachprinzip")



Beweis des Pumping-Lemma

... **Bew.**:

q := erster doppelt erreichter Zustand

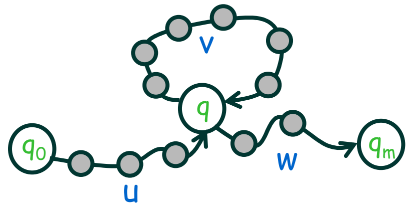
u := Wort, das zum ersten Mal (von q_0) nach q führt

v := Wort, das (von q) einmal nach q führt

("erste Wiederholung" von q)

w := Rest von z

Dann ist:



- $z=uvw$
- $|uv|\leq n$
- $|v|\geq 1$
- $uv^iw\in L$ für alle $i\geq 0$

Anwendung des Pumping-Lemmas

Idee: L regulär → Pumping Lemma gilt

also

falls der Nachweis gelingt, dass das Pumping Lemma für L nicht gilt, kann man folgern, dass L nicht regulär ist

Pumping Lemma:

$\exists n\geq 0\ \forall z\in L$ mit $|z|\geq n$:

$\exists u,v,w\in\Sigma^*$ mit $z=uvw$, $|uv|\leq n$ und $|v|\geq 1$

$\forall i\geq 0 : uv^iw\in L$

Anwendung des Pumping-Lemmas

Idee: L regulär → Pumping Lemma gilt

also

falls der Nachweis gelingt, dass das Pumping Lemma für L nicht gilt, kann man folgern, dass L nicht regulär ist

"Negiertes" Pumping Lemma:

$\forall n \geq 0 \; \exists z \in L$ mit $|z| \geq n$:

$\forall u,v,w \in \Sigma^*$ mit $z=uvw$, $|uv| \leq n$ und $|v| \geq 1$

$\exists i > 0 : \neg(uv^iw \in L)$

Nachweis der Nichtregularität von Sprachen

- $L_1 = \{0^n1^n \mid n \geq 0\}$

$\forall n \geq 0$ gilt: für $z=0^n1^n \in L$ ist $|z| \geq n$ und

$\forall u,v,w \in \Sigma^*$ mit $z=uvw$, $|uv| \leq n$ und $|v| \geq 1$

ist $uv^2w \in L^c$

(v enthält nur das Symbol 0)

Pumping Lemma gilt nicht für L_1 → L_1 nicht regulär

Nachweis der Nichtregularität von Sprachen

- $L_2 = \{ww^{rev} | w\epsilon\{0,1\}^*\} \cup \{w0w^{rev} | w\epsilon\{0,1\}^*\} \cup \{w1w^{rev} | w\epsilon\{0,1\}^*\}$

  (Palindrome über $\{0,1\}$)

  $\forall n\geq 0$ gilt: für $z=0^n1^n0^n\epsilon L$ ist $|z|\geq n$ und

  $\forall u,v,w\epsilon\Sigma^*$ mit $z=uvw$, $|uv|\leq n$ und $|v|\geq 1$

  ist $uv^2w\epsilon L^c$

  (v enthält nur das Symbol 0)

  Pumping Lemma gilt nicht für L_2 → L_2 nicht regulär

Nachweis der Nichtregularität von Sprachen

- $L_3 = \{0^{n^2} \mid n\geq 0\}$

  $\forall n\geq 0$ gilt: für $z=0^{n^2}$ ist $|z|\geq n$ und

  $\forall u,v,w\epsilon\Sigma^*$ mit $z=uvw$, $|uv|\leq n$ und $|v|\geq 1$

  ist $uv^2w\epsilon L^c$

  wegen $1 \leq |v| \leq |uv| \leq n$ ist

  $$n^2 = |z| = |uvw| < |uv^2w| = |z| + |v| \leq n^2+n < (n+1)^2$$

  Pumping Lemma gilt nicht für L_3 → L_3 nicht regulär

Bemerkungen zum Pumping-Lemma

- Pumping Lemma ist keine Charakterisierung der regulären Sprachen *("nur notwendig")*

  L regulär → Pumping-Lemma gilt

- es gibt nicht-reguläre Sprachen, die die Bedingung aus dem Pumping Lemma erfüllen

  $L = \{c^m a^n b^n \mid m,n \geq 0\} \cup \{a,b\}^*$