Übersicht

Die Myhill–Nerode–Relation $\equiv_L$

Definition 2.6.1

Es sei $L \subseteq \Sigma^*$.

Definiere $\equiv_L \subseteq \Sigma^* \times \Sigma^*$ als

$$u \equiv_L v \iff (u\,w \in L \Leftrightarrow v\,w \in L \text{ für alle } w \in \Sigma^*).$$

Der Index einer Äquivalenzrelation ist die Anzahl ihrer Äquivalenzklassen.

Interessanter Fall: $\equiv_L$ hat endlichen Index.

Beispiel 1

Es sei $L = 0^*1^*$.

- $001 \equiv_L 0111$
- $010 \not\equiv_L 0111$, denn $010 \notin L$, $0111 \in L$.
- $00 \not\equiv_L 00001$, denn $000 \in L$, $000010 \notin L$.

Wieviele Äquivalenzklassen hat $\equiv_L$?

Drei:

1. 0^*
2. 0^*1^+
3. $(0+1)^*\,1\,0\,(0+1)^*$

Beispiel 2

Was ist der Index von $\equiv_L$ für diese Sprachen?

1. $L = \{0, 1\}^*$
2. $L = \{\, a^p \mid p \text{ ist eine Primzahl} \,\}$
3. $L = \varnothing$
4. $L = \{\, w \in \{a, b, c\}^* \mid |w| \text{ ist Vielfaches von } 7 \,\}$
5. $L = \{3, 3., 3.1, 3.14, 3.141, 3.1415, 3.14159, \ldots\}$
6. $L = \{\, a^n b^n \mid n \geq 0 \,\}$

Lemma 2.6.2

$L \subseteq \Sigma^*$ *regulär* $\implies$ $\equiv_L$ *hat endlichen Index.*

Beweis.

1. L regulär.
   Daher $L = L(M)$ für einen DFA $M = (Q, \Sigma, \delta, q_0, F)$.
2. Definiere $u \sim v \iff \hat{\delta}(q_0, u) = \hat{\delta}(q_0, v)$.
3. $u \sim v \Rightarrow u \equiv_L v$, denn $u\,w \in L \iff v\,w \in L$ falls $\hat{\delta}(q_0, u) = \hat{\delta}(q_0, v)$.
4. Also hat $\sim$ mindestens so viele Äquivalenzklassen wie $\equiv_L$.
5. $\sim$ hat aber endlichen Index.

$\square$

Lemma 2.6.3

$L \subseteq \Sigma^*$ regulär $\quad \Longleftarrow \quad \equiv_L$ hat endlichen Index.

Beweis.

1. $L \subseteq \Sigma^*$ und Index von $\equiv_L$ sei endlich.
2. Konstruiere $M = (Q, \Sigma, \delta, [\epsilon]_{\equiv_L}, F)$ mit
   - $Q = \{ [w]_{\equiv_L} \mid w \in \Sigma^* \}$
   - $\delta \colon Q \times \Sigma \to Q$ mit $\delta([w]_{\equiv_L}, a) = [w\,a]_{\equiv_L}$
   - $F = \{ [w]_{\equiv_L} \mid w \in L \}$
3. Q endlich, da Index von $\equiv_L$ endlich.
4. δ wohldefiniert, da $[u]_{\equiv_L} = [v]_{\equiv_L} \Rightarrow [u\,a]_{\equiv_L} = [v\,a]_{\equiv_L}$
5. $L(M) = L$, da $\hat{\delta}([\epsilon]_{\equiv_L}, w) = [w]_{\equiv_L}$.

$\square$

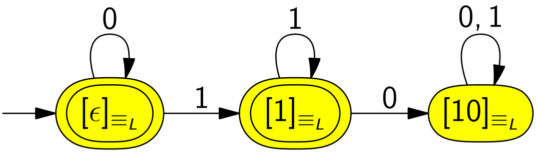
Beispiel

Es sei $L = 0^*1^*$.

$\equiv_L$ hat die Äquivalenzklassen

1. $[\epsilon]_{\equiv_L} = 0^*$,
2. $[1]_{\equiv_L} = 0^*1^+$ und
3. $[10]_{\equiv_L} = (0+1)^*\,1\,0\,(0+1)^*$.

Der Myhill–Nerode–Automat aus Lemma 2.6.3:



Der Satz von Myhill–Nerode

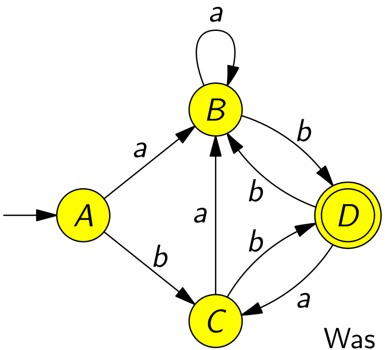
Satz 2.6.4

1. *$L \subseteq \Sigma^*$ ist genau dann regulär, wenn $\equiv_L$ endlichen Index hat.*
2. *M ein DFA $\Longrightarrow \sim_M$ ist eine Verfeinerung von $\equiv_{L(M)}$.*
3. *Es gibt zu jeder regulären Sprache $L \subseteq \Sigma^*$ einen bis auf Isomorphie eindeutigen minimalen DFA $M = (Q, \Sigma, \delta, q_0, F)$ mit $L = L(M)$.*

Beweis.

1. Folgt aus Lemma 2.6.2 und Lemma 2.6.3.
2. Beweis von Lemma 2.6.2: $u \sim v \Rightarrow u \equiv_L v$.
3. *Beweisskizze*:
   Da $\sim$ eine Verfeinerung von $\equiv_L$ ist, muss $\sim = \equiv_L$ gelten, wenn ihre Indexe gleich sind. □

Beispiel

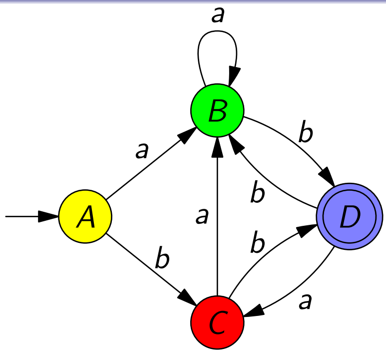


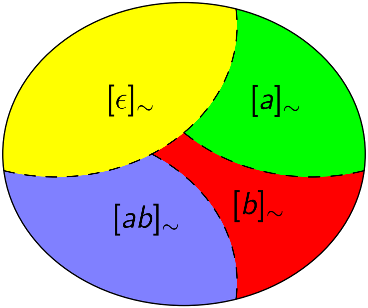
Was sind die Äquivalenzklassen von $\sim$?

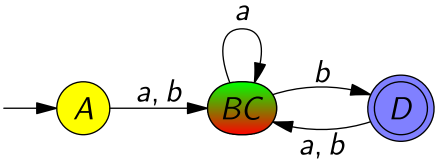
Natürlich $[\epsilon]_\sim$, $[a]_\sim$, $[b]_\sim$ und $[ab]_\sim$.

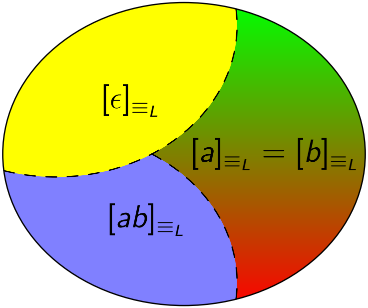
Was sind die Äquivalenzklassen von $\equiv_{L(M)}$?

Es sind $[\epsilon]_\sim$, $[a]_\sim \cup [b]_\sim$ und $[ab]_\sim$.



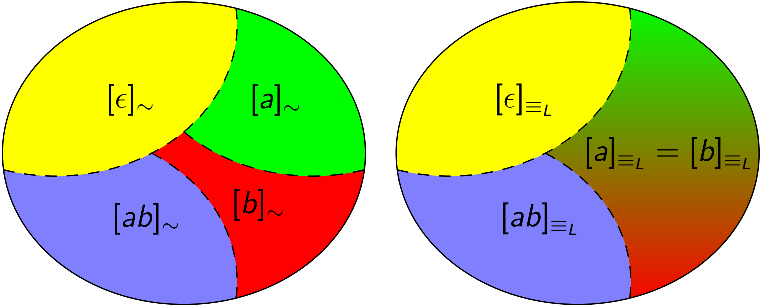






$\sim$ ist Verfeinerung von $\equiv_L$.

Eindeutigkeit des minimalen DFA



- Jede Äquivalenzklasse von $\equiv_L$ ist Vereinigung von Äquivalenzklassen von $\sim$.
- Jede Äquivalenzklasse von $\sim$ ist eindeutig einem erreichbaren Zustand zugeordnet.
- Haben $\sim$ und $\equiv_L$ den gleichen Index, dann sind sie gleich.

Definition 2.6.5

Es seien DFAs gegeben:

- $M = (Q, \Sigma, \delta, q_0, F)$
- $M' = (Q', \Sigma, \delta', q_0', F')$

Eine Abbildung $h\colon Q \to Q'$ *mit*

1. $h(q) \in F' \iff q \in F$
2. $h(q_0) = q_0'$
3. $h(\delta(q, a)) = \delta'(h(q), a)$

heißt Homomorphismus von M *nach* M'*.*

Ist h *bijektiv, dann ist es ein Isomorphismus.*

Beweis von Satz 2.6.4 (3)

Es $M = (Q, \Sigma, \delta, q_0, F)$ ein DFA mit $L(M) = L$.

$\sim$ definiert als: $u \sim v \iff \hat{\delta}(q_0, u) = \hat{\delta}(q_0, v)$.

Sei $M' = (Q', \Sigma, \delta', [\epsilon]_\sim, F')$ mit

- $Q' = \Sigma^*/\sim$ (Äquivalenzklassen von $\sim$)
- $\delta'([w]_\sim, a) = [w\,a]_\sim$
- $F' = \{\, [w]_\sim \mid w \in L(M) \,\}$

M und M' sind isomorph.

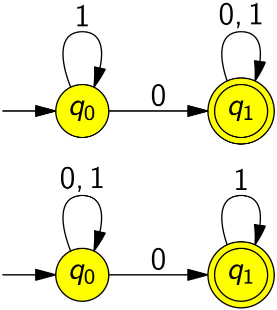
$h(q) = \{\, w \in \Sigma^* \mid \hat{\delta}(q_0, w) = q \,\}$ ist ein Isomorphismus.

- M' hängt nur von L und $\sim$ ab.
- $\sim \, = \, \equiv_L$, falls M minimal.
- Also hängt M' *nur* von L ab (der Myhill–Nerode–DFA).

Folgerung: Alle minimalen Automaten sind isomorph.

Frage:

Sind auch kleinste NFAs isomorph?



Gegenbeispiel! Beide akzeptieren $(0+1)^*0(0+1)^*$.

Die Eindeutigkeit des minimalen DFAs ist etwas besonderes!

Andere Konsequenz des Satzes von Myhill–Nerode:

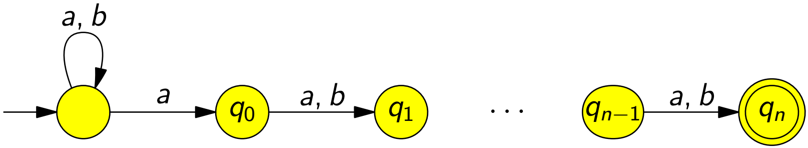
> *Die Anzahl der Zustände des minimalen Automaten für L ist der Index von $\equiv_L$.*

Zwei wichtige Anwendungen:

1. Untere Schranke für die Anzahl der Zustände eines DFA, der L akzeptiert.
2. Beweis, dass eine Sprache nicht regulär ist.

(1) Untere Schranke für Zustände eines DFA

Es sei $L = (a+b)^* a (a+b)^n$ mit $n \in \mathbf{N}$.
NFA für L:



Es sind $n+2$ Zustände.

Wähle $K = (a+b)^n$.
Behauptung: Falls $u, v \in K$ mit $u \neq v$, dann $u \not\equiv_L v$.
Beweis:
o.B.d.A. $u = w\,a\,u'$, $v = w\,b\,v'$. Dann $ua^{n-|u'|} \in L$, $va^{n-|u'|} \notin L$.

Also hat $\equiv_L$ mindestens $|K| = 2^n$ viele Äquivalenzklassen.
Jeder DFA der L akzeptiert, hat mindestens 2^n Zustände.

(2) Beweis, dass Sprache nicht regulär ist

Es sei $L = \{ a^n b^n \mid n \geq 0 \}$.

Wähle $K = a^*$.

Wieder gilt:
$u, v \in K$, $u \neq v$, dann $u \not\equiv_L v$.

Denn: $a^i\, b^i \in L$, $a^j\, b^i \notin L$, falls $a^i \neq a^j$.

Index von $\equiv_L$ ist mindestens $|K| = \infty$.

Wäre L regulär, dann hätte der minimale DFA mindestens $|K|$ Zustände.

Das beweist, dass L nicht regulär ist.

(2) Beweis, dass Sprache nicht regulär ist

Es sei $L = \{ a^p \mid p \text{ ist Primzahl } \}$.

Vorüberlegung:
Es seien $p_1 < p_2$ zwei Primzahlen und $d = p_2 - p_1$, d.h. $d \geq 1$.
Betrachte $p_1 + n * d$ mit $1 \leq n \leq p_1$.

Behauptung:
Es gibt ein n mit

- $1 \leq n \leq p_1$
- $p_1 + n * d$ ist prim
- $p_1 + (n+1) * d = p_2 + n * d$ ist nicht prim

Wähle $K = L$.
Es seien $a^{p_1}, a^{p_2} \in K$ mit $p_1 < p_2$.

Dann ist $a^{p_1} a^{n*d} \in L$ und $a^{p_2} a^{n*d} \notin L$.

Also hat $\equiv_L$ unendlichen Index, d.h. L ist nicht regulär.