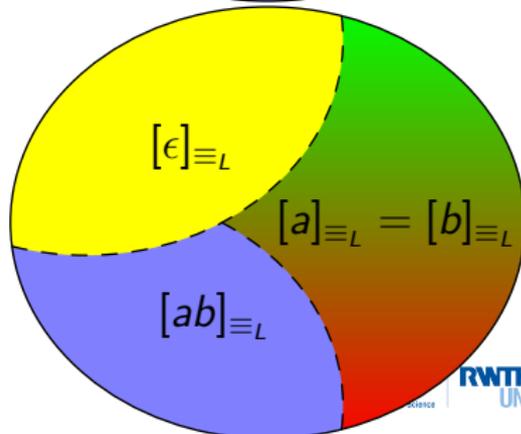
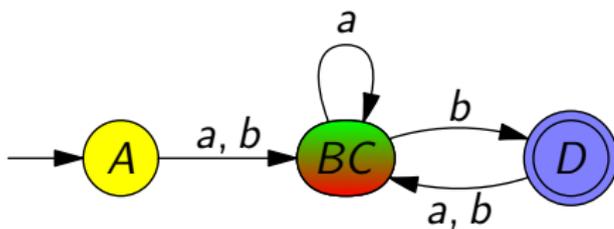
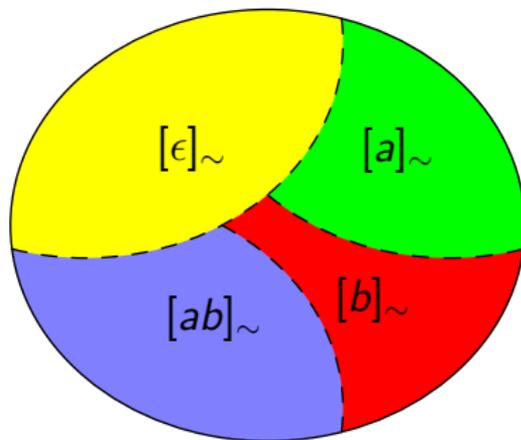
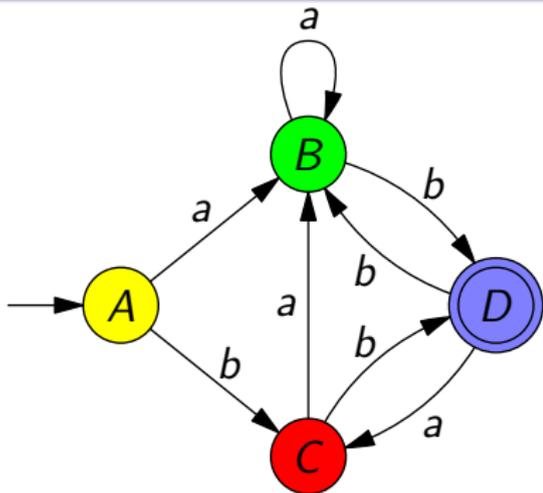


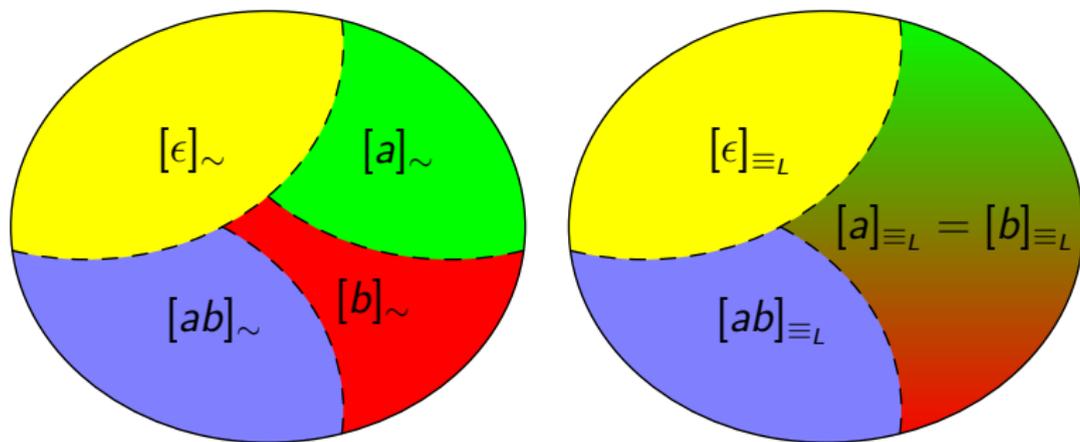
Reguläre Sprachen

Minimale DFAs und der Satz von Myhill–Nerode



\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 Zustand zugeordnet.
- Haben \sim und \equiv_L den gleichen Index, dann sind sie gleich.

Definition

Es seien DFAs gegeben:

- $M = (Q, \Sigma, \delta, q_0, F)$
- $M' = (Q', \Sigma, \delta', q'_0, F')$

Eine Abbildung $h: Q \rightarrow Q'$ 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*.

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

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

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

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

M und M' sind isomorph.

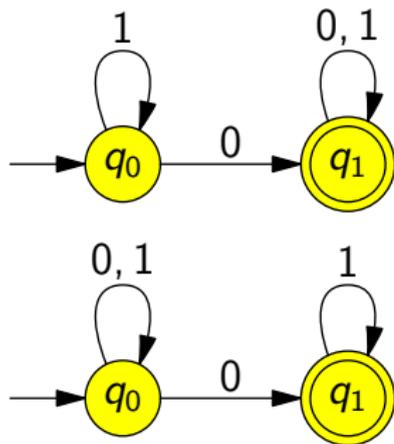
$h: q \mapsto \{ w \in \Sigma^* \mid \hat{\delta}(q_0, w) = q \}$ ist ein Isomorphismus.

Folgerung: Alle minimalen Automaten sind isomorph:

- 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)

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:

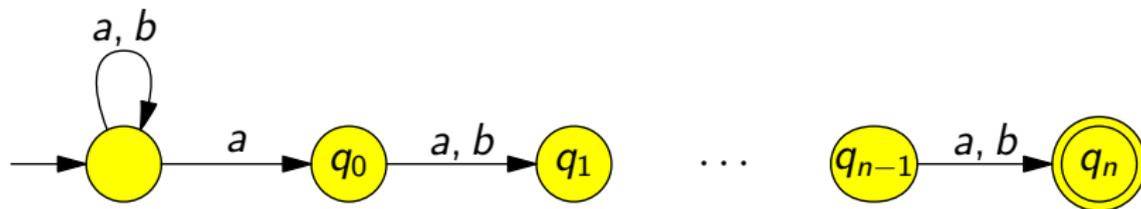
Die Anzahl der Zustände des minimalen Automaten für L ist der Index von \equiv_L .

Zwei wichtige Anwendungen:

- 1 Untere Schranken für die Anzahl der Zustände.
- 2 Beweis, daß eine Sprache nicht regulär ist.

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

NFA für L :



Es sind $n + 2$ Zustände.

Wähle $N = (a + b)^n$.

Behauptung: Falls $u, v \in N$ mit $u \neq v$, dann $u \not\equiv_L v$.

Beweis:

o.B.d.A. $u = wau'$, $v = wbv'$. Dann $ua^{n-|u'|} \in L$, $va^{n-|u'|} \notin L$.

Also hat \equiv_L mindestens $|N| = 2^n$ viele Äquivalenzklassen.

Jeder DFA der L akzeptiert, hat mindestens 2^n Zustände.

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

Wähle $N = a^*$.

Wieder gilt:

$u, v \in N$, $u \neq v$, 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 $|N| = \infty$.

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

Das beweist, daß L nicht regulär ist.

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

Vorüberlegung:

Es seien $p_1 < p_2$ zwei Primzahlen und $d = p_2 - p_1$.

Betrachte $p_1 + nd$ mit $1 \leq n \leq p_1$.

Behauptung:

Es gibt ein n mit

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

Wähle $N = L$.

Es seien $a^{p_1}, a^{p_2} \in N$ mit $p_1 < p_2$.

Dann ist $a^{p_1} a^{nd} \in L$ und $a^{p_2} a^{nd} \notin L$.

Also hat \equiv_L unendlichen Index.

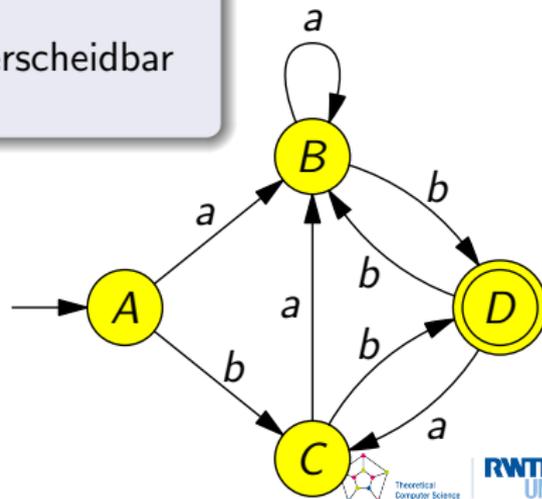
Minimierung von DFAs

Definition

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

$q_1, q_2 \in Q$ sind *unterscheidbar*, falls

- 1 $q_1 \in F, q_2 \notin F$ oder
- 2 $q_1 \notin F, q_2 \in F$ oder
- 3 $\delta(q_1, a)$ und $\delta(q_2, a)$ sind unterscheidbar für ein $a \in \Sigma$.



Lemma

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

$q_1, q_2 \in Q$ unterscheidbar genau dann wenn:

Es gibt $w \in \Sigma^*$ mit

- $\hat{\delta}(q_1, w) \in F, \hat{\delta}(q_2, w) \notin F$ oder
- $\hat{\delta}(q_2, w) \in F, \hat{\delta}(q_1, w) \notin F$.

Beweis.

„ \Leftarrow “ Induktion über $|w|$:

$|w| = 0$: Dann $\hat{\delta}(q_i, w) = q_i$ und damit $q_1 \in F$ und $q_2 \notin F$ oder umgekehrt.

$|w| > 0$: Es sei $w = au$ mit $u \in \Sigma^*$ und $a \in \Sigma$.

$q'_1 = \delta(q_1, a)$ und $q'_2 = \delta(q_2, a)$ unterscheidbar nach I.V.:

$\hat{\delta}(q'_1, u) \in F, \hat{\delta}(q'_2, u) \notin F$ oder umgekehrt.

„ \Rightarrow “ Übungsaufgabe!



Markierungsalgorithmus

$M := \emptyset;$

for((q, p) in $Q \times Q$) {

if($p \in F \wedge q \notin F$) $M := M \cup \{(q, p)\};$

if($p \notin F \wedge q \in F$) $M := M \cup \{(q, p)\};$

}

do {

$M_{\text{old}} := M;$

for((q, p) $\in Q \times Q$)

for($a \in \Sigma$)

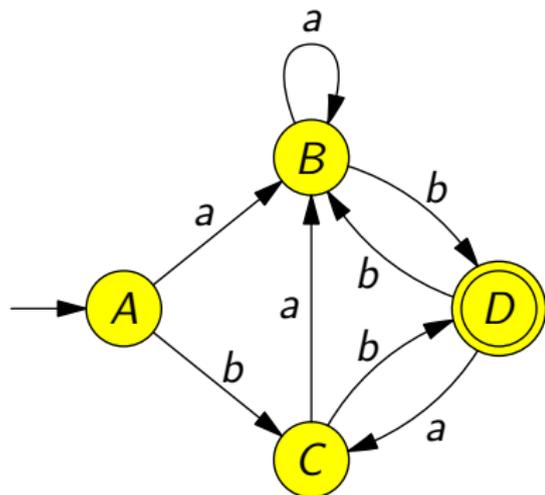
if(($\delta(q, a), \delta(p, a)$) $\in M$) $M := M \cup \{(q, p)\};$

while($M_{\text{old}} \neq M$);

Laufzeit: Sehr grobe Abschätzung $O(|Q|^4)$

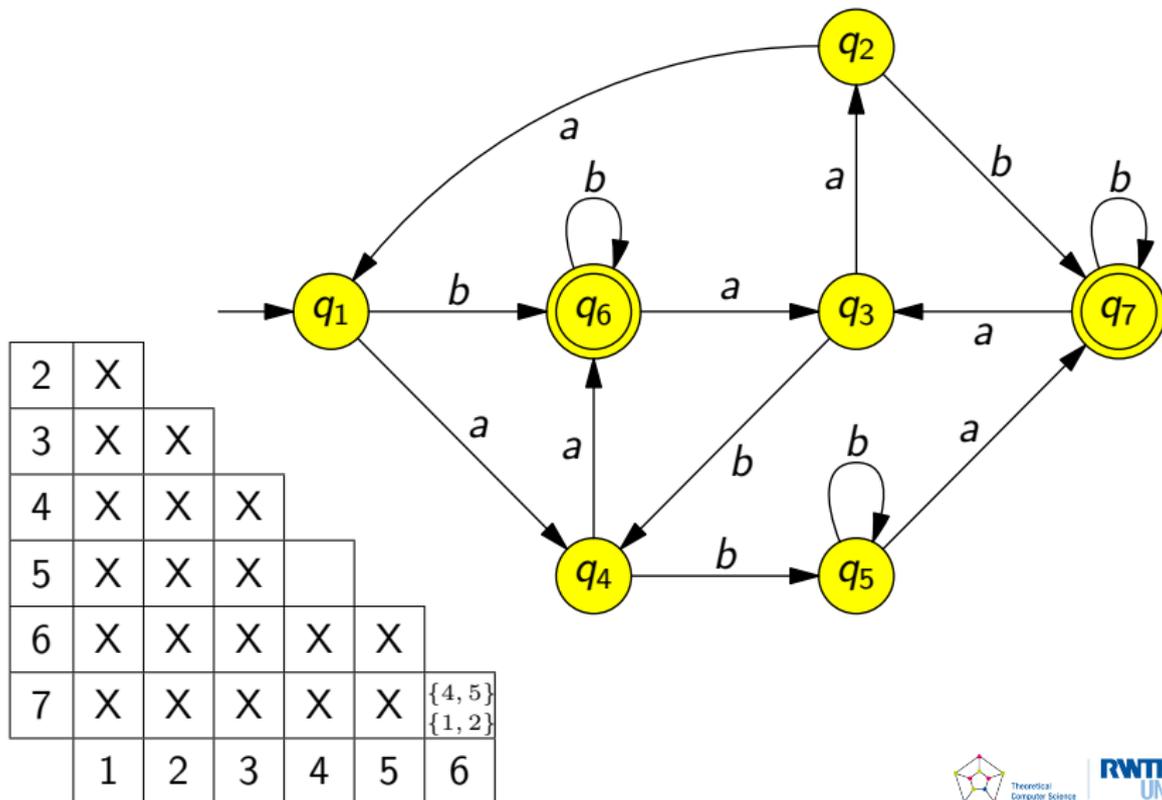
- Maximal $|Q|^2$ Wiederholungen (jedesmal ein Paar markiert)
- Jede Wiederholung in $O(|Q|^2)$ Schritten

Beispiel 1



M	A	B	C	D
A		×	×	×
B	×			×
C	×			×
D	×	×	×	

Beispiel 2



Schnellerer Markierungsalgorithmus

```
M := ∅;
for((q, p) ∈ Q × Q) {
  if(p ∈ F ∧ q ∉ F) M := M ∪ {(q, p)};
  if(p ∉ F ∧ q ∈ F) M := M ∪ {(q, p)};
}
for((q, p) ∈ Q × Q)
  for(a ∈ Σ) {
    q' := δ(q, a); p' := δ(p, a);
    if((q', p') ∈ M) {
      M := M ∪ {(q, p)};
      Füge rekursiv L(q, p) in M ein
    }
    else L(q', p') := L(q', p') ∪ {(q, p)};
  }
}
```

Laufzeit: Mit amortisierter Analyse $O(|Q|^2)$

Konstruktion des minimalen DFA

Lemma

Es sei $M = (Q, \Sigma, \delta, q_0, F)$ ein DFA und $p, q \in Q$.

Weiter sei $p = \hat{\delta}(q_0, u)$ und $q = \hat{\delta}(q_0, v)$.

Falls $u \equiv_{L(M)} v$, dann sind u und v nicht unterscheidbar.

Beweis.

Andernfalls gäbe es w mit $\hat{\delta}(p, w) \in F \Leftrightarrow \hat{\delta}(q, w) \notin F$.

Dann gilt auch $\hat{\delta}(q_0, uw) \in F \Leftrightarrow \hat{\delta}(q_0, vw) \notin F$.

Damit ist $uw \in L(M) \Leftrightarrow vw \notin L(M)$.

Das ist ein Widerspruch zu $u \equiv_{L(M)} v$. □

Konstruktion des minimalen DFA

Theorem

Verschmelzen wir die nicht unterscheidbaren Zustände eines DFA, erhalten wir den zugehörigen minimalen DFA.

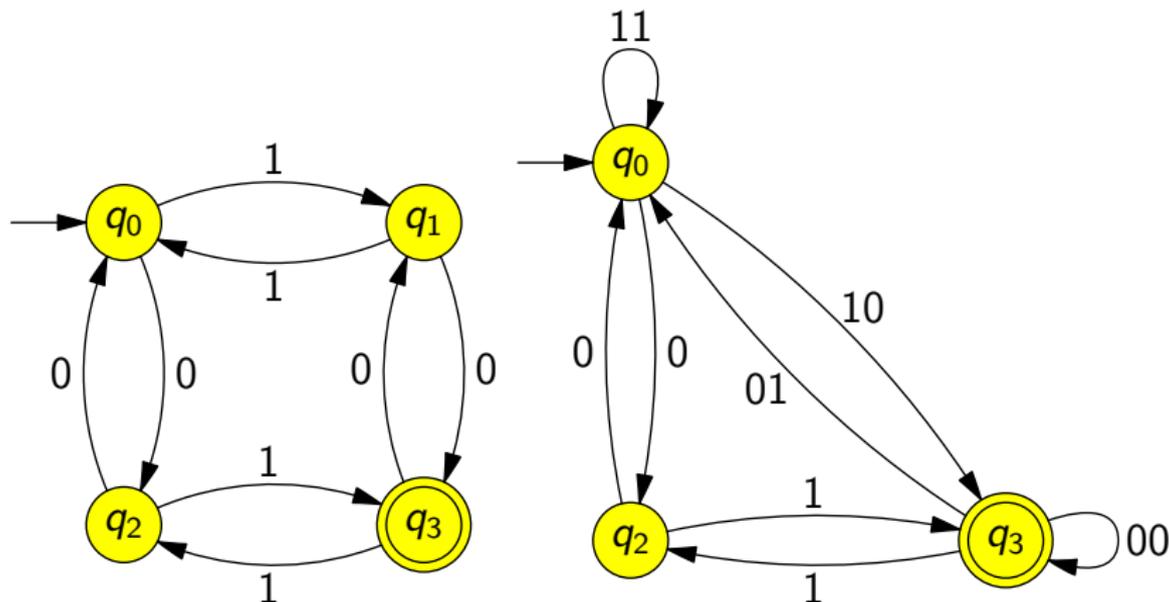
Beweis.

Letztes Lemma:

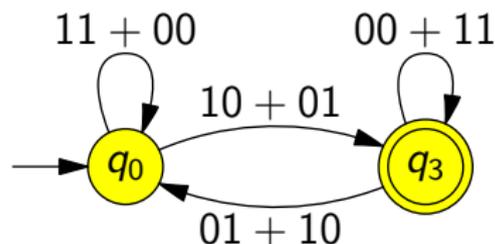
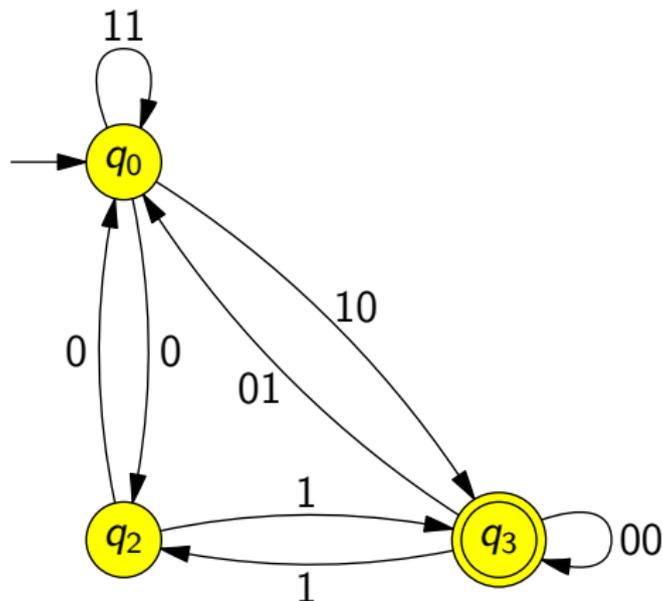
Falls keine unterscheidbaren Zustände, dann $\sim = \equiv_{L(M)}$.

Satz von Myhill–Nerode: Der Automat ist minimal. □

Eliminierung von Zuständen



Statt Symbolen: Reguläre Ausdrücke auf Übergängen!



Regulärer Ausdruck:

$$(11 + 00)^*(10 + 01)(00 + 11 + (01 + 10)(11 + 00)^*(10 + 01))^*$$

Sehr kurz!