# Theoretische Informatik
# WS 2024/2025

Fachhochschule Dortmund
University of Applied Sciences and Arts
we focus on students



## Kapitel 2

## *Reguläre Sprachen*

## 2.7

## *Grenzen regulärer Sprachen*

Prof. Dr. Robert Preis
Fachbereich Informatik
Fachhochschule Dortmund
Robert.Preis@fh-dortmund.de

# Grenzen der regulären Sprachen



*Welche Sprachen gehören zu den regulären Sprachen?*

*Und wie beweise ich das?*

*Und welche gehören nicht zu den regulären Sprachen?*

*Und wie beweise ich das?*

# Abschlusseigenschaften von Operationen über Mengen



Abschlusseigenschaften erleichtern den Umgang.

*Die natürlichen Zahlen sind abgeschlossen über „+“*, d.h.

- Immer wenn ich zwei natürliche Zahlen addiere, dann kommt immer wieder ein natürliche Zahl heraus.
- Z.B. 198+365 = 563 ist eine natürliche Zahl

Abschlusseigenschaften sind nicht selbstverständlich, denn:

*Die natürlichen Zahlen sind NICHT abgeschlossen über „-“*, d.h.

- Wenn ich zwei natürliche Zahlen subtrahiere, kommt nicht immer eine natürliche Zahl heraus.
- Z.B. 198-365 = -167 ist keine natürliche Zahl

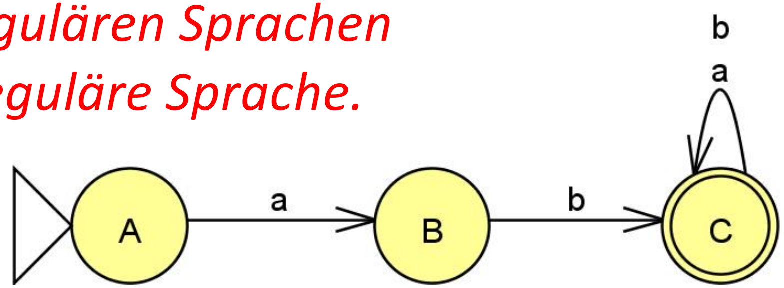
Einige Operationen auf regulären Sprachen führen immer zu regulären Sprachen.
*Welche Operationen sind dies?*

# Abschlusseigenschaften von regulären Sprachen: Beispiel

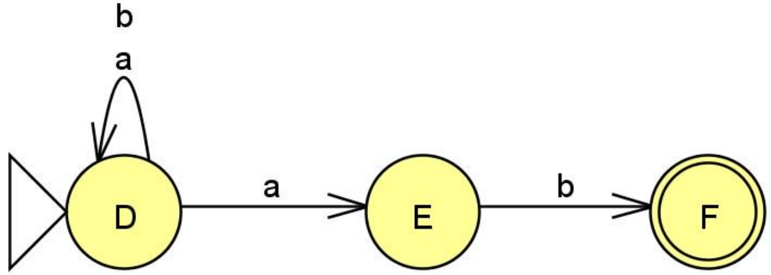


*Die Vereinigung von zwei regulären Sprachen ergibt immer wieder eine reguläre Sprache.*

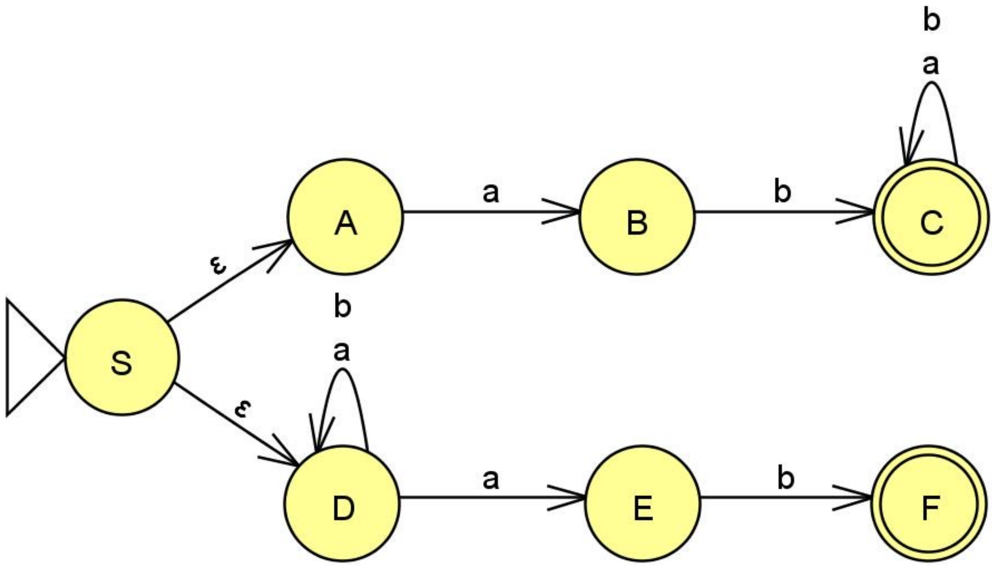
*$L_1 = \{w \in \{a,b\}^* \mid w$ fängt mit ‚ab' an$\}$ ist regulär:*




*$L_2 = \{w \in \{a,b\}^* \mid w$ hört mit ‚ab' auf$\}$ ist regulär:*




*$L = L_1 \cup L_2$ ist regulär:*



# Abschlusseigenschaften von regulären Sprachen



- $L_1 \cup L_2$ Die Vereinigung zweier regulärer Sprachen ist regulär.
- $L_1 \cap L_2$ Der Durchschnitt zweier regulärer Sprachen ist regulär.
- $L_1 \setminus L_2$ Die Differenz zweier regulärer Sprachen ist regulär.
- $L^c$ Das Komplement einer regulären Sprache ist regulär.
- $L^R$ Die Spiegelung einer regulären Sprache ist regulär.
- $L_1 \circ L_2$ Die Verkettung zweier regulärer Sprachen ist regulär.
- $L^*$ Die Hülle einer regulären Sprache ist regulär.

Wie beweist man das?

Wir können dabei alle Modelle von regulären Sprachen (rechtslineare Grammatik, DEA, NEA, ε-NEA, regulärer Ausdruck) verwenden:

- Durch Reguläre Ausdrücke wird bewiesen: $\cup$, $\circ$, $*$
- Durch Automaten wird bewiesen: $^c$, $^R$
- Durch Mengengesetze wird bewiesen: $\cap$, $\setminus$

# Beweise für Abschluss unter Vereinigung, Verkettung und Hülle



Beweisführung mit **regulären Ausdrücken**:

1. Seien $L_1$, $L_2$ regulär, d.h. es gibt reguläre Ausdrücke
   - $E_1$ mit $L_1 = L(E_1)$
   - $E_2$ mit $L_2 = L(E_2)$.
2. Dann ist nach der Definition von Regulären Ausdrücken auch
   - $L_1 \cup L_2 \quad = L(E_1) \cup L(E_2) = L(E_1 | E_2)$ regulär
   - $L_1 \circ L_2 \quad = L(E_1) \circ L(E_2) = L(E_1 \circ E_2)$ regulär
   - $L_1{}^* \quad = L(E_1)^* = L(E_1{}^*)$ regulär

Beispiele: $E_1 = (a|b)$ $\qquad E_2 = b^*$

- Vereinigung: $L((a|b)) \cup L(b^*) \quad = L((a|b) \mid b^*)$.
- Konkatenation: $L((a|b)) \circ L(b^*) \quad = L((a|b) \circ b^*)$.
- Hülle: $L((a|b))^* \quad = L((a|b)^*)$.

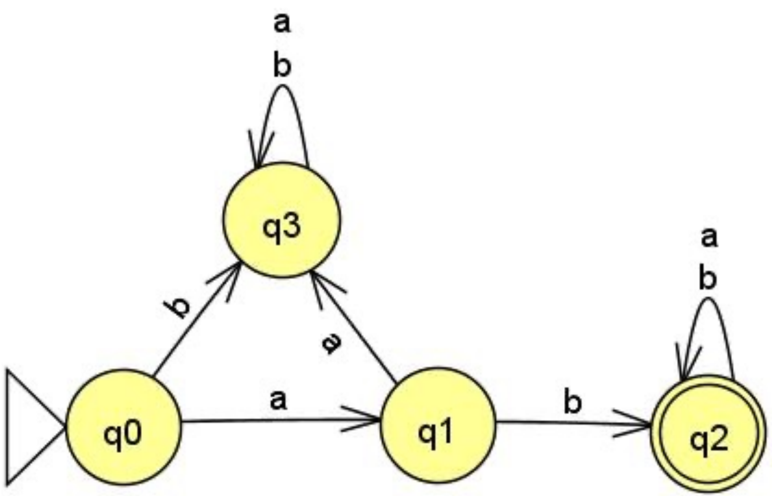
# Beweise für Abschluss unter Komplement, Durchschnitt, Differenz

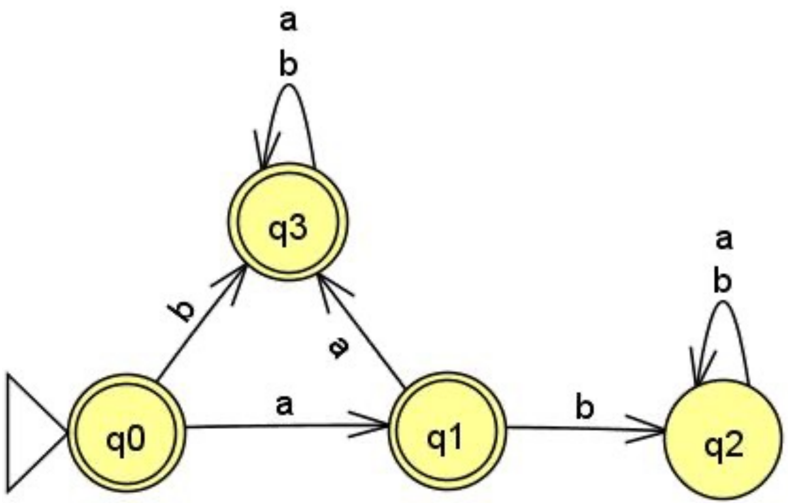


Beweisführung mit **endlichen Automaten**:

1. Sei L eine reguläre Sprache, d.h. es gibt einen DEA A mit L(A) = L.
2. Konstruiere aus A einen Automaten B, bei dem die Endzustände mit den Nicht-Endzuständen vertauscht sind.
   Der Automat B akzeptiert $L^c$, d.h. $L^c$ ist eine reguläre Sprache.

Beispiel: Wörter, die mit ab anfangen:




Einfache Beweise durch **Mengengesetze**:

$L_1$, $L_2$ regulär $\Rightarrow$ $L_1 \cap L_2 = (L_1^c \cup L_2^c)^c$ regulär

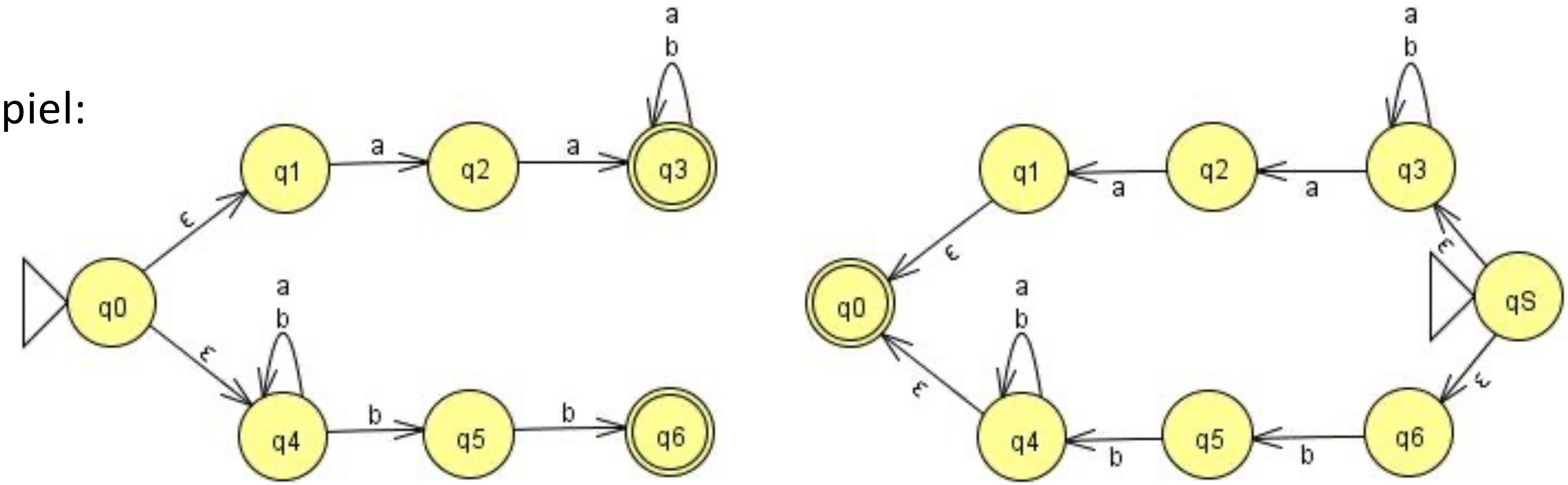
$L_1$, $L_2$ regulär $\Rightarrow$ $L_1 \backslash L_2 = L_1 \cap L_2^c$ regulär

# Beweis für Abschluss unter Spiegelung

Beweisführung mit **endlichen Automaten**:

1. Sei L eine reguläre Sprache, d.h. es gibt Automaten A mit L(A) = L.
2. Konstruiere aus A einen Automaten B, bei dem
   - alle Pfeile umgekehrt sind
   - $q_0$ wird neuer und einziger Endzustand.
   - $q_S$ wird als neuer Zustand der Startzustand mit ε-Übergängen zu allen bisherigen Endzuständen.
3. Der so konstruierte Automat akzeptiert $L^R$, d.h. $L^R$ ist reguläre Sprache.

Beispiel:



# Ist eine Sprache regulär?

Ist ((((((((((((((((((()))))))))))))))))

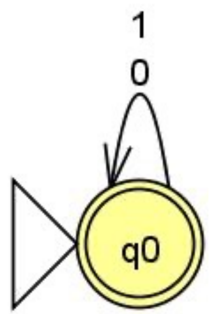
ein gültiger Klammerausdruck?

Sind die korrekten Klammerausdrücke überhaupt eine reguläre Sprache?

Ist überhaupt $L = \{ w \in \{(,)\}^* \mid w = (^n\ )^n,\ n \in \mathbf{N}_+ \}$ regulär?

Ich kann auch fragen:

Ist überhaupt $L = \{ w \in \{0,1\}^* \mid w = 0^n 1^n,\ n \in \mathbf{N}_+ \}$ regulär?




Was passiert, wenn ich versuche, dafür einen DEA zu erstellen?

Dieser Automat akzeptiert zwar alle Wörter aus L, aber auch weitere Wörter!

Es gibt 2 Möglichkeiten, um zu zeigen, dass eine Sprache nicht regulär ist:

1. Durch Abschlusseigenschaften (einfach, man muss aber eine Sprache als Referenz haben, von der man weiß, dass sie nicht regulär ist)
2. Zeigen, dass das Pumping Lemma nicht gilt (was für jede reguläre Sprache gelten muss)

# Zeige, dass eine Sprache NICHT regulär ist: (1) Abschlusseigenschaft



*Zeige durch Widerspruch,*
*dass wenn eine Sprache L regulär wäre,*
*wegen einer Abschlusseigenschaft*
*auch eine andere Sprache regulär sein müsste,*
*von der aber bekannt ist, dass sie nicht regulär ist.*

Wir setzen als bekannt voraus (Beweis durch Pumping Lemma möglich):

$L_1 = \{w \in \{0,1\}^* \mid w=0^n1^n, n \in \mathbf{N_0}\}$ ist keine reguläre Sprache.

*Ist $L_2 = \{w \in \{0,1\}^* \mid w$ hat gleich viele 0'en und 1'en $\}$ eine reguläre Sprache?*

Annahme: $L_2$ ist eine reguläre Sprache.

- $L_3 = \{w \in \{0,1\}^* \mid w=0^n1^m, n,m \in \mathbf{N_0}\}$ ist eine reguläre Sprache (einfach zu zeigen).
- Wegen der Abschlusseigenschaft des Durchschnitts wäre auch die Sprache $L_2 \cap L_3 = \{w \in \{0,1\}^* \mid w=0^n1^n, n \in \mathbf{N_0}\} = L_1$ regulär … Widerspruch !

=> D.h. die Annahme ist falsch und es gilt: $L_2$ ist keine reguläre Sprache!

# Zeige, dass eine Sprache NICHT regulär ist: (2) das Pumping Lemma muss gelten

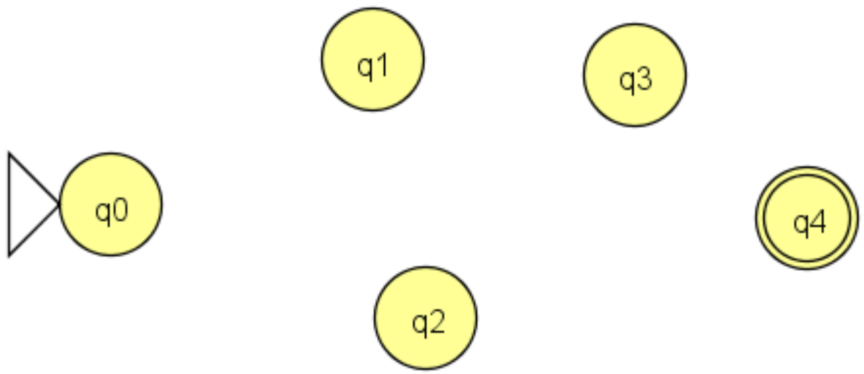


Endliche Sprachen sind immer regulär.

Wir betrachten jetzt unendliche Sprachen. Darin muss es beliebig lange Wörter geben (sonst wäre die Sprache endlich).

Ein DEA hat aber nur endlich viele Zustände!

Beispiel:




$|Q|$ = 5 Zustände

(Übergänge sind nicht dargestellt)

Wie lang sind Wörter ohne Kreise/Schleifen? ***Max. Länge |Q|-1 !***

Wie können sehr lange Wörter verarbeitet werden? ***Über Kreise oder Schleifen !***

***Jedes Wort der Länge >= |Q|***

***muss einen Kreis (oder eine Schleife) enthalten !***

***Nach spätestens |Q| Schritten ist der Kreis gebildet !***

# Idee: Durchlaufen des Kreises beliebig oft

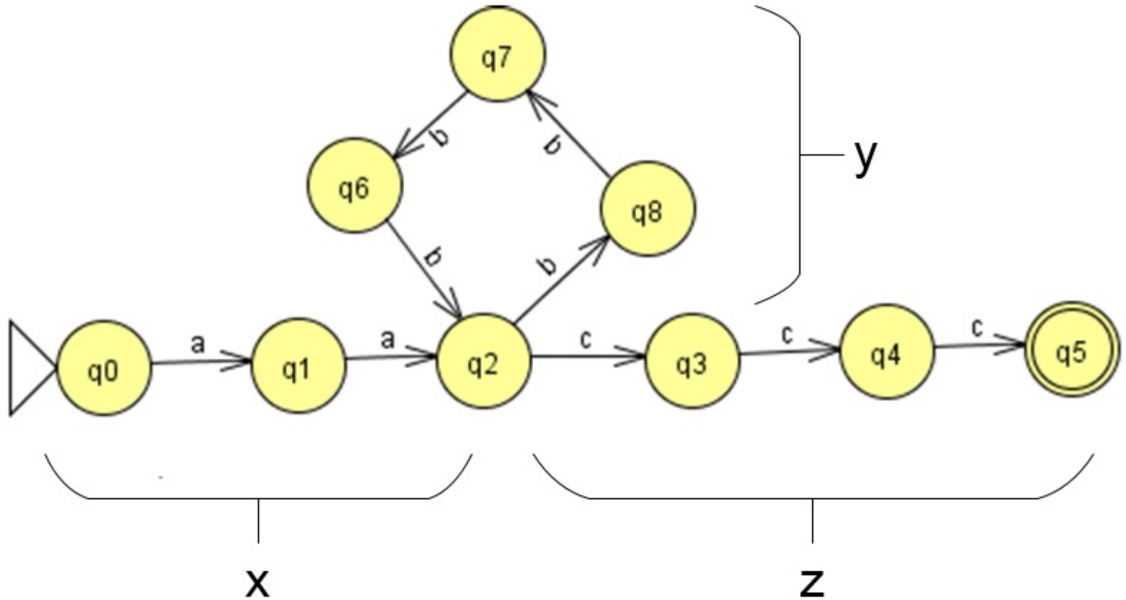
Durchlaufe ich beim Abarbeiten des Wortes einen Kreis, dann kann ich den Kreis auch mehrfach (k-mal) durchlaufen. Diese Wörter müssen dann auch in der Sprache sein.

Ich kann den Kreis auch weglassen. Dieses Wort muss dann auch in der Sprache sein.

Betrachte das Wort

w = xyz , w ∈ L :

(Nicht alle Zustände und
Übergänge sind dargestellt!)



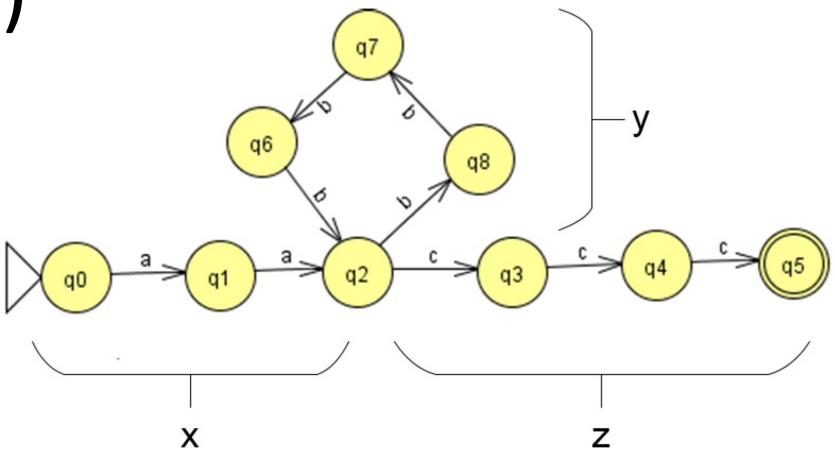

Dann sind auch alle Wörter $w = xy^kz$ in der Sprache L,

d.h. $w = xy^kz \in L(A)$, $k \geq 0$, im Beispiel: aaccc, aabbbbccc, aabbbbbbbbccc,... ∈ L(A)

d.h. das Wort w ist aufpumpbar.

# Das Pumping Lemma formal (gilt für alle regulären Sprachen)






**Pumping Lemma**:

Für jede reguläre Sprache L gibt es eine Zahl p∈**N** (p ist die Pumpinglänge oder die Pumpingzahl), so dass jedes Wort w∈L mit |w|≥p zerlegt werden kann in w = xyz mit den Eigenschaften (auf Grundlage eines DEA für L):

1. $y \neq \varepsilon$ (der Kreis ist nicht leer)
2. $|xy| \leq p$ (der Kreis ist nach höchstens p Schritten gebildet)
3. $xy^kz \in L$ für alle $k \in \mathbf{N}_0$ (egal, wie oft man den Kreis durchläuft, es werden auch weitere Wörter mit demselben Anfang x und demselben Ende z akzeptiert)

*Gibt es immer ein p?*

*Ja, einfach p=„Anzahl der Zustände im entsprechenden DEA" (im Bsp. p = 9) !*

# Wenn das Pumping Lemma nicht gilt, dann ist die Sprache auch nicht regulär...



Behauptung: $L = \{0^m 1^m \mid m \in \mathbf{N}\}$ ist nicht regulär.

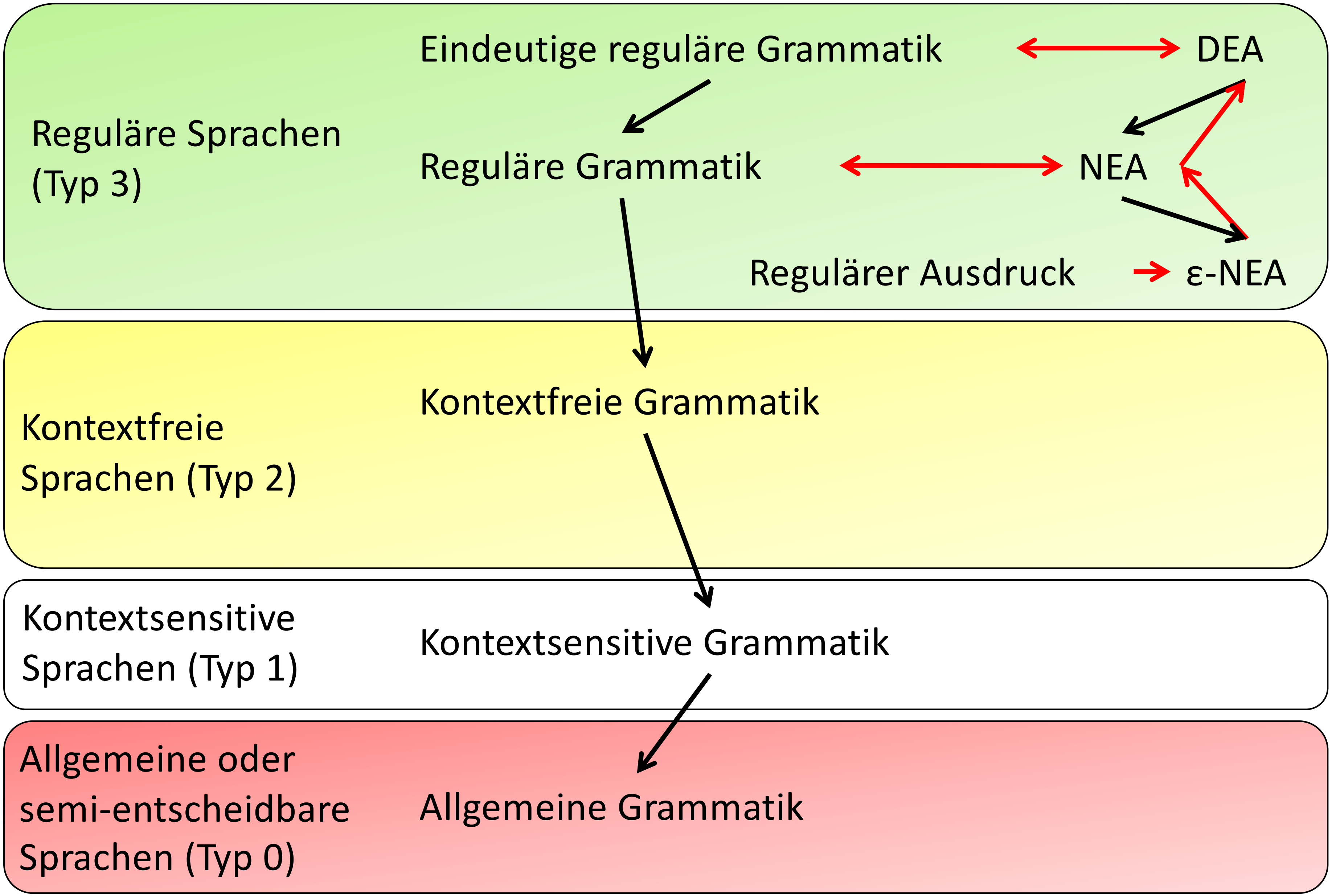
Beweis durch Widerspruch im Pumping Lemma:

1. **Annahme: L ist regulär**
2. Dann gibt es einen DEA A mit L(A) = L.
3. A hat mindestens einen Zustand (Startzustand) und somit hat A $n \in \mathbf{N}$ Zustände. Wir wissen nicht genau wie groß n ist, nur dass es eine natürliche Zahl ist...
4. A akzeptiert alle Wörter aus L und wir wählen daraus geschickt das Wort $\mathbf{w=0^n1^n}$ aus, ohne genau zu wissen, wie groß n ist.
5. Es gilt das Pumping Lemma und $|0^n1^n|>=n$. D.h. man kann w zerlegen in w=xyz mit $y \neq \varepsilon$ und $|xy| \leq n$ (y ist Kreis, der nach spätestens n Übergängen auftritt).
6. Dann bestehen x und y nur aus 0'en und es gilt $x=0^i$, $y=0^j$ und $z=0^{n-i-j}1^n$ für ein $j \geq 1$ und $i+j \leq n$.
7. Laut Pumping Lemma muss dann auch $xy^kz \in L$ für jedes $k \in \mathbf{N}$ sein. Wir wählen k=0, d.h. $xy^0z=xz=0^i0^{n-i-j}1^n=0^{n-j}1^n \in L$. Daraus folgt $j=0$.
8. $j \geq 1$ und $j=0$ ist ein Widerspruch. **Deshalb kann L nicht regulär sein.**

# Überblick Reguläre Sprachen

| | Typ-3 | Typ-2 | Typ-1 | Typ-0 |
|---|---|---|---|---|
| Frage 1: L={} ? | √ | | | |
| Frage 2: $w \in L$ ? | √ | | | |
| Frage 3: $L_1=L_2$ ? | √ | | | |
| Frage 4: minimal ? | √ | | | |
| $L_1 \cup L_2$ | √ | | | |
| $L_1 \cap L_2$ | √ | | | |
| $L_1 \setminus L_2$ | √ | | | |
| $L^c$ | √ | | | |
| $L_1 \circ L_2$ | √ | | | |
| $L^*$ | √ | | | |

# Überblick der Transformationen



# Zusammenfassung



- Reguläre Sprachen haben einige Abschlusseigenschaften:
  - $\cup$, $\cap$, $\setminus$, $^c$, $^R$, $\circ$, $*$
  - Man kann damit (in manchen Fällen) beweisen, dass eine Sprache regulär ist oder nicht.
- Pumping Lemma:
  - Eine reguläre Sprache hat die Eigenschaft, dass wenn lange Wörter akzeptiert werden, man auf einen Kreis stößt und man durch beliebig häufiges Durchlaufen des Kreises unendlich viele Wörter der Sprache generieren kann.
  - Man kann damit ggf. beweisen, dass eine Sprache nicht regulär ist (und zwar indirekt, mit einem „Widerspruchsbeweis“, weil ein einziges Gegenbeispiel ausreicht, um zu zeigen, dass das Pumping-Lemma für die Sprache nicht gilt.)