# Determinisierung



Ziel: für jeden NEA $A$ gibt es einen äquivalenten DEA $A_D$

Idee:

- zu jedem Zeitpunkt Menge von möglichen Zuständen $S \subseteq Q$
- $|Q|$ endlich $\rightsquigarrow$ endlich viele Teilmengen
- für $S \subseteq Q$ und Eingabe $a$ ist die Menge der möglichen Folgezustände eindeutig

$\rightsquigarrow$ DEA $A_D$, dessen Zustände Teilmengen von $Q$ sind

---

Prinzipiell wäre es vorstellbar, dass der Nichtdeterminismus es erlaubt, Sprachen zu erkennen, die ein deterministischer Automat nicht erkennen kann. Wir werden nun zeigen, dass dies nicht so ist:

Zu jedem NFA gbt es einen DFA; der die gleiche Sprache erkennt. Die Idee der Konstruktion ist oben skizziert und auf der nächsten Folie präzise ausgeführt.

Endliche Automaten
Äquivalenz der Automatenmodelle
Teilmengen-konstruktion
Exponentieller Blow-up
$\epsilon$-Übergänge
Reguläre Ausdrücke
Pumping Lemma

# Teilmengenkonstruktion

Sei $A_N = (Q_N, \Sigma, \delta_N, q_0, F_N)$ ein NEA.

Definiere daraus einen DEA $A_D$ mit

- Zuständen $Q_D = 2^{Q_N}$.
- für $S \in Q_D$, also $S \subseteq Q_N$, und $a \in \Sigma$ ist

$$\delta_D(S, a) = \bigcup_{q \in S} \delta_N(q, a)$$

- Anfangszustand ist $\{q_0\}$
- Endzustände in $F_D$ sind Mengen $S$ mit

$$S \cap F_N \neq \emptyset$$

Die Zustände von $A_D$ sind *Teilmengen* der Zustandsmenge von $A_N$. Idee ist, dass diese die möglichen Zustände darstellt, in denen sich $A_N$ zu einem Zeitpunkt befinden kann.

Zur Definition von $\delta_D$: Argument ist ein Zustand von $A_D$, also eine Menge $S$ von Zuständen von $A_N$, und ein Symbol $a$. für jeden Zustand $q \in S$ gibt $\delta_N(q, a)$ eine Menge von möglichen Folgezuständen. Die Menge der möglichen Folgezustände nach einem der Zustände in $S$ ist also die Vereinigung der $\delta_N(q, a)$ für $q \in S$. Vergleiche dies auch mit der Definition von $\hat{\delta}$ für NEA.

Der NEA $A_N$ akzeptiert ein Wort $w$, wenn sich unter den möglichen Zuständen nach Lesen von $w$ ein Endzustand aus $F$ befindet, daher sind die Endzustände von $A_D$ diejenigen Mengen, die mindestens einen Endzustand enthalten. Vergleiche auch hierzu die Akzeptanzbedingung des NEA.

# Beweis der Äquivalenz



## Theorem

*Sei $A_N = (Q_N, \Sigma, \delta_N, q_0, F_N)$ ein NEA,*

*und $A_D = (Q_D, \Sigma, \delta_D, \{q_0\}, F_D)$ der mit der Teilmengenkonstruktion gewonnene DEA.*

*Dann ist $L(A_D) = L(A_N)$.*

## Lemma

*Seien $A_N$ und $A_D$ wie oben. Es gilt für alle $w \in \Sigma^*$:*

$$\hat{\delta}_D(\{q_0\}, w) = \hat{\delta}_N(q_0, w)$$

Beweis des Theorems:

$$\begin{array}{llll}
w \in L(A_N) & \text{gdw.} & \hat{\delta}_N(q_0, w) \cap F_N \neq \emptyset & \text{Akzeptanzbedingung von } A_N \\
& \text{gdw.} & \hat{\delta}_D(\{q_0\}, w) \cap F_N \neq \emptyset & \text{nach dem Lemma} \\
& \text{gdw.} & \hat{\delta}_D(\{q_0\}, w) \in F_D & \text{Definition von } F_D \\
& \text{gdw.} & w \in L(A_D) & \text{Akzeptanzbedingung von } A_D
\end{array}$$

Beweis des Lemmas durch Induktion nach $|w|$:
Induktionsanfang: für $w = \epsilon$ ist laut Definition der Funktionen $\hat{\delta}$:

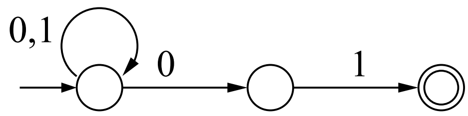
$$\hat{\delta}_D(\{q_0\}, \epsilon) = \{q_0\} = \hat{\delta}_N(q_0, \epsilon)$$

Induktionsschritt: für $w = va$ gilt

$$\begin{array}{lll}
\hat{\delta}_D(\{q_0\}, va) = & \delta_D(\hat{\delta}_D(\{q_0\}, v), a) & \text{nach Definition von } \hat{\delta}_D \\
= & \delta_D(\hat{\delta}_N(q_0, v), a) & \text{nach Induktionshypothese} \\
= & \displaystyle\bigcup_{q \in \hat{\delta}_N(q_0, v)} \delta_N(q, a) & \text{nach Definition von } \delta_D \\
= & \hat{\delta}_N(q_0, va) & \text{nach Definition von } \hat{\delta}_N
\end{array}$$

# Beispiel

| | 0 | 1 |
|---|---|---|
| $\rightarrow q_0$ | $\{q_0, q_1\}$ | $\{q_0\}$ |
| $q_1$ | $\emptyset$ | $\{q_2\}$ |
| $* \; q_2$ | $\emptyset$ | $\emptyset$ |




| | | 0 | 1 |
|---|---|---|---|
| | $\emptyset$ | $\emptyset$ | $\emptyset$ |
| $\rightarrow$ | $\{q_0\}$ | $\{q_0, q_1\}$ | $\{q_0\}$ |
| | $\{q_1\}$ | $\emptyset$ | $\{q_2\}$ |
| $*$ | $\{q_2\}$ | $\emptyset$ | $\emptyset$ |
| | $\{q_0, q_1\}$ | $\{q_0, q_1\}$ | $\{q_0, q_2\}$ |
| $*$ | $\{q_0, q_2\}$ | $\{q_0, q_1\}$ | $\{q_0\}$ |
| $*$ | $\{q_1, q_2\}$ | $\emptyset$ | $\{q_2\}$ |
| $*$ | $\{q_0, q_1, q_2\}$ | $\{q_0, q_1\}$ | $\{q_0, q_2\}$ |

Der DEA in der unteren Tabelle entseht durch die Teilmengenkonstruktion wie in der Definition aus dem NEA oben. Für jede der 8 Teilmengen der Zustandsmenge gibt es eine Zeile, also einen Zustand im DEA.

# Erreichbarkeit

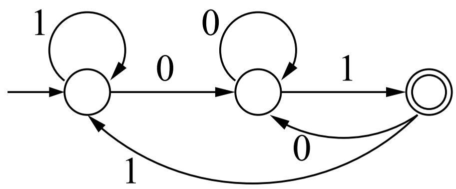


Automat $D$ hat $|Q_D| = 2^{|Q_N|}$ Zustände.

Aber: Brauchen nur erreichbare Zustände:

| | | 0 | 1 |
|---|---|---|---|
| | $\emptyset$ | $\emptyset$ | $\emptyset$ |
| $\rightarrow$ | $\{q_0\}$ | $\{q_0, q_1\}$ | $\{q_0\}$ |
| | $\{q_1\}$ | $\emptyset$ | $\{q_2\}$ |
| $*$ | $\{q_2\}$ | $\emptyset$ | $\emptyset$ |
| | $\{q_0, q_1\}$ | $\{q_0, q_1\}$ | $\{q_0, q_2\}$ |
| $*$ | $\{q_0, q_2\}$ | $\{q_0, q_1\}$ | $\{q_0\}$ |
| $*$ | $\{q_1, q_2\}$ | $\emptyset$ | $\{q_2\}$ |
| $*$ | $\{q_0, q_1, q_2\}$ | $\{q_0, q_1\}$ | $\{q_0, q_2\}$ |

$\rightsquigarrow$ Äquivalenter DEA:




Bei der Anwendung der Teilmengenkonstruktion wird üblicherweise gleich nur der erreichbare Teil konstruiert. Man beginnt mit dem Anfangszustand $\{q_0\}$, bildet die Übergänge dazu, trägt dann die auftetenden Zustandsmengen in die Tabelle ein, und fährt so fort, bis keine neuen Mengen mehr auftreten. Beispiele dazu gibt es in den Übungen.

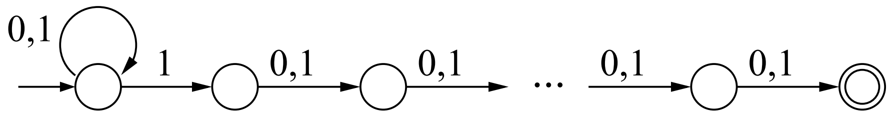
Das Beispiel zeigt, dass in manchen Fällen der erreichbare Teil deutlich weniger als $2^{|Q|}$ Mengen enthält. Das Beispiel auf der nächsten Folie zeigt, dass es jedoch Sprachen gibt, für die jeder DEA exponentiell grösser sein muss als ein NEA.

# Exponentieller Blow-up



Sei $L_n := \{w\,;\, w = u1v \text{ mit } |v| = n-1\}$.

$L_n$ wird erkannt durch NEA $A_n$ mit $n+1$ Zuständen:




## Theorem

*Jeder DEA, der $L_n$ erkennt, hat mindestens $2^n$ Zustände.*

Die Sprache $L_n$ enthält genau die Wörter, bei denen an der $n$-ten Stelle von hinten eine 1 steht. Offensichtlich akzeptiert der NEA $A_n$ diese Sprache.

Der informelle Grund, warum ein DEA für die Sprache $L_n$ groß sein muss, ist dass er sich immer die letzten $n$ Symbole merken muss. Dies wird formalisiert in dem folgenden Beweis:

Beweis des Theorems:

Sei $A = (Q, \{0,1\}, \delta, q_0, F)$ ein DEA mit $L(A) = L_n$.

Die Annahme $|Q| < 2^n$ wird zum Widerspruch führen, also ist $|Q| \geq 2^n$.

Definiere für jedes $w \in \{0,1\}^n$ den Zustand $q_w = \hat{\delta}(q_0, w)$.
Ist $|Q| < 2^n$, so gibt es verschiedene $v, w \in \{0,1\}^n$ mit $q_v = q_w$.

Da $v \neq w$, müssen sie sich einer Stelle $i \leq n$ unterscheiden. ObdA sei das $i$-te Symbol in v eine 0 und das $i$-te Symbol in $w$ eine 1.

Dann ist aber $v0^{i-1} \notin L_n$ und $w0^{(i-1)} \in L_n$. Andererseits ist aber

$$\hat{\delta}(q_0, v0^{i-1}) = \hat{\delta}(q_v, 0^{i-1}) = \hat{\delta}(q_w, 0^{i-1}) = \hat{\delta}(q_0, w0^{i-1})$$

da $q_v = q_w$ ist. Also sind $v0^{i-1}$ und $w0^{(i-1)}$ entweder beide in $L(A)$, oder keines von beiden, im Widerspruch zu $L(A) = L_n$. $\stackrel{?}{=}$