Lektion 13

In dieser Lektion lernen Sie eine weitere Form endlicher Automaten kennen, die *nichtdeterministischen endlichen Automaten*. Nichtdeterministisch bedeutet dabei, dass es mehrere Möglichkeiten geben kann, in welchen Zustand ein Automat im nächsten Schritt wechselt. Im ersten Moment mag diese Form von Automaten eigenartig erscheinen, es zeigt sich aber, dass das Konzept "Nichtdeterminismus" bei der Verwendung und Verarbeitung von Sprachen durchaus hilfreich sein kann.

Kap. 8 Nichtdeterministische endliche Automaten

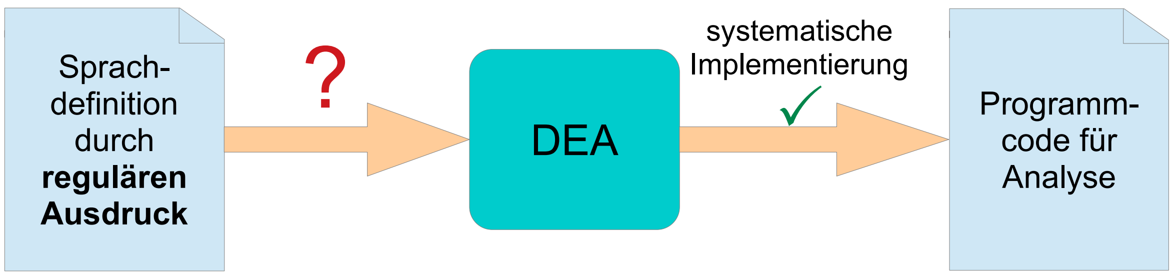
Inhalt

- Definition nichtdeterministischer endlicher Automaten (ε-NEA)
- Zusammenhang zwischen nichtdeterministischen und deterministischen endlichen Automaten
- Zusammenhang zwischen regulären Ausdrücken und endlichen Automaten
- Anwendung endlicher Automaten in Scanner-Generatoren

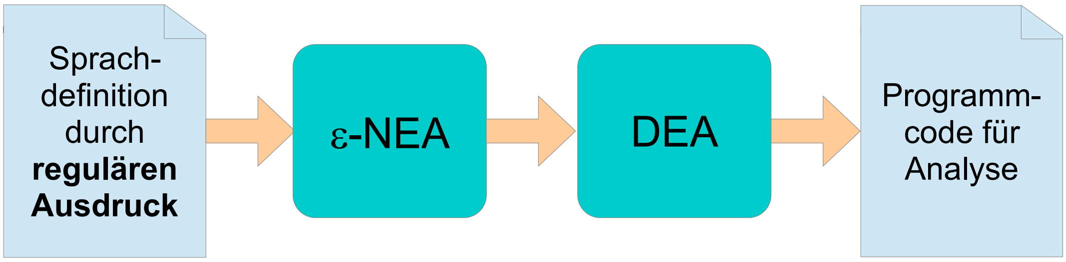
8.1 Motivation

Im vorigen Kapitel haben Sie bereits gesehen, wie *deterministische* Automaten verwendet werden können, um zu prüfen, ob ein Wort zu einer Sprache gehört. Sie haben auch gesehen, dass DEAs einfach, systematisch und effizient implementiert werden können.

Nichtdeterministische Automaten werden eine entscheidende Rolle spielen, wenn es darum geht, zu einem gegebenen regulären Ausdruck einen DEA zu bilden, der die gleiche Sprache akzeptiert, d.h. mit dem geprüft werden kann, ob ein Wort zur Sprache gehört.



Die direkte "Übersetzung" eines regulären Ausdrucks in einen DEA ist allerdings sehr schwierig. Mit Hilfe eines Zwischenschritts, bei dem zunächst ein sog. nichtdeterministischer Automat (ε-NEA) gebildet wird, der dann nachfolgend in einen deterministischen umgewandelt werden kann, ist es aber relativ einfach möglich.



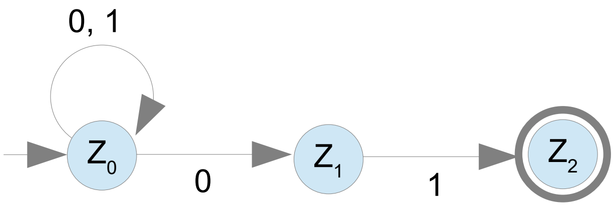
In den folgenden Abschnitten lernen Sie zunächst, wie nichtdeterministische Automaten aufgebaut sind, welche Möglichkeiten zur Beschreibung von Sprachen sie bieten und auch, wie sie in äquivalente deterministische Automaten umgewandelt werden können.

8.2 Nichtdeterministische endliche Automaten mit ε-Übergängen

Vor der formalen Definition betrachten wir zunächst zwei Beispielautomaten, die die wesentlichen Möglichkeiten nichtdeterministischer Automaten zeigen.

Beispiel 8.1 - Nichtdeterminismus in Automaten

Betrachten Sie folgenden endlichen Automaten. Welche Zeichenfolgen akzeptiert der Automat? Was fällt Ihnen auf?

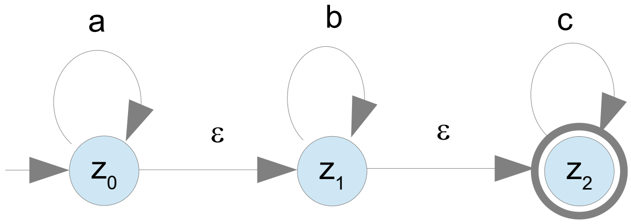


Eine naheliegende Antwort wäre die, dass der Automat alle Zeichenfolgen über Alphabet $\Sigma = \{0,1\}$ akzeptiert, die mit 01 enden (als regulärer Ausdruck ließe sich das mit (0|1)*01 beschreiben).

Bei genauerer Betrachtung sieht man, dass es kein deterministischer Automat ist, da es in Zustand z_0 für das Zeichen 0 zwei mögliche Nachfolgezustände z_0 und z_1 gibt. Der Automat ist deshalb ein *nichtdeterministischer* endlicher Automat.

Der folgende Automat mit Eingabealphabet $\Sigma = \{a,b,c\}$ nutzt eine weitere Eigenschaft von nichtdeterministischen endlichen Automaten.

Beispiel 8.2 - Automat mit ε-Übergängen



Es ist möglich, dass der Automat über "spontane" Zustandsübergänge, die mit ε gekennzeichnet sind, von einen Zustand in einen anderen Zustand wechselt, ohne dass ein Symbol des Eingabeworts verarbeitet wird.

Die genaue formale Definition, wann ein Wort von einem nichtdeterministischen Automat akzeptiert wird, folgt später. Es wird dann (nicht allzu überraschend) so sein, dass der Automat aus Beispiel 8.2 alle Wörter akzeptiert, die mit beliebig vielen Zeichen *a* beginnen, danach können beliebig viele *b* folgen und am Ende können noch beliebig viele *c* kommen, d.h. alle Wörter der Sprache des regulären Ausdrucks $a^*b^*c^*$.

Definition 8.3 - Nichtdeterministischer endlicher Automat (ε -NEA)

Ein **nichtdeterministischer endlicher Automat mit ε-Übergängen** (**ε-NEA**) $A = (Z, \Sigma, \Delta, z_0, E)$ ist definiert durch:

Z	endlicher Zustandsmenge
Σ	Eingabealphabet
$z_0 \in Z$	Startzustand
$\Delta \subseteq Z \times (\Sigma \cup \{\varepsilon\}) \times Z$	***Zustandsübergangsrelation***
$E \subseteq Z$	Menge von akzeptierenden Zuständen (Endzuständen)

Der einzige Unterschied zur Definition von DEAs ist der, dass statt einer Zustandsübergangs*funktion* eine Zustandsübergangs*relation* verwendet wird.

Zustandsübergangsrelation Δ - Erläuterung und Notation

Die Möglichkeit des Automaten, sich nichtdeterministisch verhalten zu können, wird mathematisch dadurch ausgedrückt, dass die möglichen Zustandsübergänge nicht durch eine *Funktion*, sondern durch eine *Relation* beschrieben werden. Zwei Zustände stehen in Relation, wenn es möglich ist, vom einen Zustand mit einem Zeichen oder auch spontan in den anderen Zustand zu wechseln. Zustandsübergänge sind also entweder mit einem Zeichen $a \in \Sigma$ verbunden oder es sind spontane Übergängen, die durch ε gekennzeichnet sind:

- Stehen *z*, *a* und *z*' in Relation, d.h. (*z*, *a*, *z*') $\in \Delta$, wobei a $\in \Sigma$, dann bedeutet das, dass es möglich ist, vom Zustand *z* in den Zustand *z*' zu wechseln, wenn das nächste Symbol im Eingabewort das Zeichen *a* ist. Wir notieren das meist folgendermaßen:

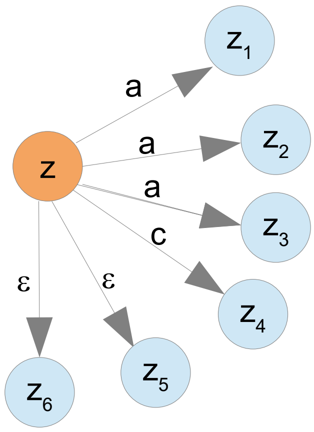
  $z \xrightarrow{a} z'$ (regulärer *a*-Übergang)

- Stehen *z*, ε und *z*' in Relation, d.h. (*z*, ε, *z*') $\in \Delta$, dann bedeutet das, dass der Automat "von sich aus" spontan von Zustand *z* in den Zustand *z*' wechseln kann, unabhängig vom nächsten Symbol des Eingabeworts. Das notieren wir auch folgendermaßen:

  $z \xrightarrow{\varepsilon} z'$ (ε-Übergang)

Anmerkungen

- *A*usgehend von einem Zustand *z* kann es bei einem ε-NEA beliebig viele Folgezustände geben, die mit den gleichen Zeichen oder über ε erreichbar sind, wie hier für einen Zustand z veranschaulicht ist (wobei Σ = {a, b, c}):



- Im Unterschied zu den DEAs ist es auch erlaubt, dass es bei einem Zustand für bestimmte Zeichen *keinen* Nachfolgezustand gibt (z.B. hat Zustand z keinen b-Übergang). Fehlerzustände wie bei DEAs sind deshalb bei ε-NEAs nicht nötig.

- ε-NEAs bieten erweiterte Beschreibungsmöglichkeiten im Vergleich zu DEAs. Jeder DEA lässt sich deshalb auch als nichtdeterministischer Automat betrachten, wobei

  (z, *a*, z') $\in \Delta$

  genau dann, wenn

  $\delta(z,a) = z'$,

  d.h. die Möglichkeit des Nichtdeterminismus und der ε-Übergänge wird dann eben nur nicht ausgenützt.

- Im Unterschied zu DEAs gibt es für ε-NEAs keine direkte, einfache Möglichkeit, sie effizient zu implementieren.

- In der Literatur finden Sie teilweise auch die Variante des **NEA,** der zwar Nichtdeterminismus in Form mehrerer Nachfolgezustände für das gleiche Zeichen erlaubt, aber keine spontanen ε-Übergänge. Wir behandeln hier aber diese Variante nicht separat.

8.2.1 Sprache eines ε-NEA

Die folgende Definition legt fest, welche Zustände mit einem gegebenen Wort erreicht werden können. Die Betonung liegt hierbei auf "können". Dies wird durch folgende auf Wörter erweiterte Zustandsübergangsrelation beschrieben.

Definition 8.4 - Übergangsrelation für Wörter

Für einen ε-NEA $A = (Z, \Sigma, \Delta, z_0, E)$ ist die **auf Wörter erweiterte Zustandsübergangsrelation**

$$\rightarrow \ \subseteq \ Z \times \Sigma^* \times Z$$

folgendermaßen definiert:

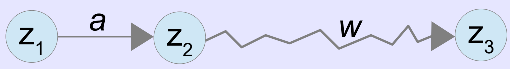
- Mit dem leeren Wort ε kann man den gleichen Zustand erreichen, d.h. für alle Zustände $z \in Z$ gilt

  $$z \xrightarrow{\varepsilon} z$$

- Gibt die Zustandsübergangsrelation Δ an, dass der Automat mit Zeichen *a* von einem Zustand z_1 nach z_2 wechseln kann und kann er mit dem Wort *w* von z_2 aus Zustand z_3 erreichen, dann kann der Automat mit Wort *aw* von z_1 aus Zustand z_3 erreichen:

  $$\text{falls } (z_1, a, z_2) \in \Delta \text{ und } z_2 \xrightarrow{w} z_3, \text{ dann } z_1 \xrightarrow{aw} z_3$$

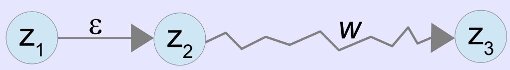
  (für alle z_1, $z_2, z_3 \in Z$, $a \in \Sigma$, $w \in \Sigma^*$)

- Gibt die Zustandsübergangsrelation Δ an, dass der Automat mit ε-Übergang von z_1 nach z_2 wechseln kann, und kann er mit dem Wort *w* von z_2 den Zustand z_3 erreichen, dann kann man mit Wort *w* auch von z_1 nach z_3 kommen:

  $$\text{falls } (z_1, \varepsilon, z_2) \in \Delta \text{ und } z_2 \xrightarrow{w} z_3, \text{ dann } z_1 \xrightarrow{w} z_3$$

  (für allez_1, $z_2, z_3 \in Z$, $w \in \Sigma^*$)

Anmerkungen

- $z \xrightarrow{w} z'$ bedeutet: von Zustand *z* aus ***kann*** mit den Zeichen des Worts *w* und ggf. davor, dazwischen oder danach liegenden ε-Übergängen der Zustand *z'*

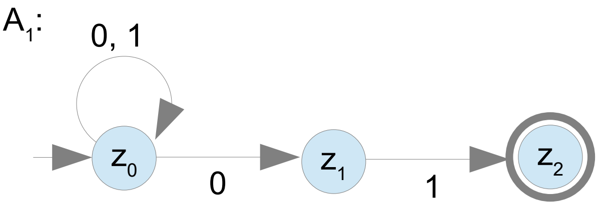
THU Technische Hochschule Ulm

erreicht werden.

- Im Unterschied zu deterministischen Automaten ist es möglich, dass mit dem gleichen Wort *w* mehrere verschiedene Zustände erreicht werden.

Beispiel 8.5 - Zustandsübergangsrelation für Wörter

(1) Gegeben ist folgender ε-NEA A_1.



Es gilt dann beispielsweise:

$z_0 \xrightarrow{0} z_0$

$z_0 \xrightarrow{0} z_1$

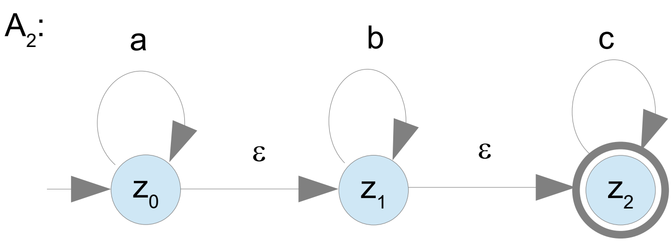
$z_0 \xrightarrow{10} z_0$

$z_0 \xrightarrow{10} z_1$

$z_0 \xrightarrow{101} z_0$

$z_0 \xrightarrow{101} z_2$

(2) A_2 sei folgender ε-NEA:



Dann gilt beispielsweise:

$z_0 \xrightarrow{\varepsilon} z_0$

$z_0 \xrightarrow{\varepsilon} z_1$

$z_0 \xrightarrow{\varepsilon} z_2$

$z_0 \xrightarrow{aaa} z_0$

$z_0 \xrightarrow{aaa} z_1$

$z_0 \xrightarrow{aaa} z_2$

$z_0 \xrightarrow{aaabb} z_1$

$z_0 \xrightarrow{aaabb} z_2$

$z_0 \xrightarrow{aaabbcccc} z_2$

Man sieht, dass mit dem gleichen Wort mehrerer Zustände erreicht werden können. Mit dem Wort *aaabb* kann beispielsweise ausgehend von Zustand z_0 sowohl z_1 als auch Zustand z_2 erreicht werden (aber nicht z_0)

Aufgabe 8.6 - Zustandsübergangsrelation für Wörter

Was trifft für den oben angegebenen Automat A_2 zu?

(1) $z_0 \xrightarrow{bb} z_0$

(2) $z_0 \xrightarrow{bb} z_1$

(3) $z_0 \xrightarrow{bb} z_2$

Die Sprache des nichtdeterministischen Automaten wird nun folgendermaßen "optimistisch" definiert:

Definition 8.7 - Sprache eines ε-NEA

Die von einem **ε-NEA** $A = (Z, \Sigma, \Delta, z_0, E)$ **akzeptierte Sprache L(*A*)** ist

$$L(A) = \{ w \in \Sigma^* \mid \text{es gibt ein } z \in E : z_0 \xrightarrow{w} z \} \quad .$$

Die akzeptierte Sprache eines ε-NEA ist die Menge aller Wörter, mit denen ausgehend von Startzustand z_0 ein Endzustand erreicht werden **kann**.

Beispiel 8.8 - Sprache eines ε-NEA

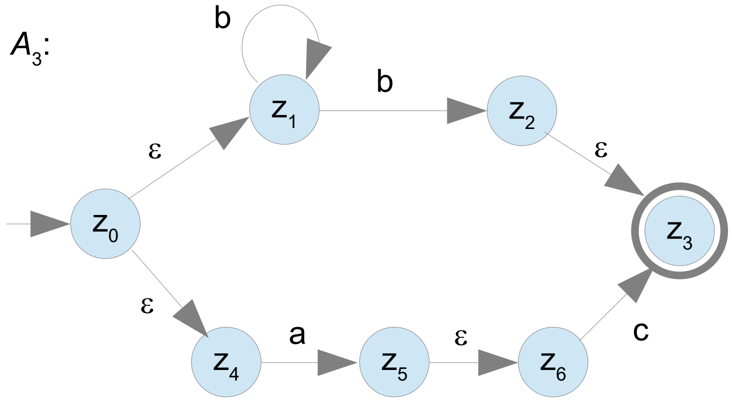
Für die Automaten aus *Beispiel 8.5 - Zustandsübergangsrelation für Wörter* gilt:

$L(A_1) = \{w01 \mid w \in \{0,1\}^* \}$ alle Wörter, die mit 01 enden.

$L(A_2) = \{a^n b^k c^j \mid n \geq 0, k \geq 0, j \geq 0 \}$ d.h. Sprache $a^*b^*c^*$

Aufgabe 8.9 - Sprache eines ε-NEA

Welche Sprache akzeptiert folgender ε-NEA A_3 über Alphabet $\Sigma = \{a, b, c\}$?



Aufgabe 8.10 - ε-NEA definieren

Definieren Sie einen ε-NEA, der alle Wörter über dem Alphabet Σ = {a, b, c} akzeptiert, die (mindestens) eines der Wörter

aba, **bab** oder **cba**

als Teilwort enthalten (an beliebiger Stelle).

8.3 Transformation von ε-NEAs in DEAs

Welche Zusammenhang besteht zwischen DEAs und ε-NEAs? Es ist klar, dass man alle Sprachen, die man mit DEAs beschreiben kann, auch mit ε-NEAs beschreiben kann. Wie sieht es aber umgekehrt aus?

- Sind nichtdeterministische Automaten mächtiger als deterministische? D.h. gibt es Sprachen, die von einem ε-NEA, aber von keinem DEA akzeptiert werden?
- Oder kann zu jedem nichtdeterministischen Automaten auch ein äquivalenter deterministische Automat konstruiert werden, d.h. einer, der die gleiche Sprache akzeptiert?

Sie werden sehen, dass die Antwort beim zweiten Punkt "ja" ist. Das bedeutet, dass ε-NEAs zwar ggf. "bequemer" sind, wenn es darum geht, Sprachen zu beschreiben, sie sind aber nicht mächtiger. D.h. damit kann man nicht mehr Sprachen beschreiben als mit DEAs.

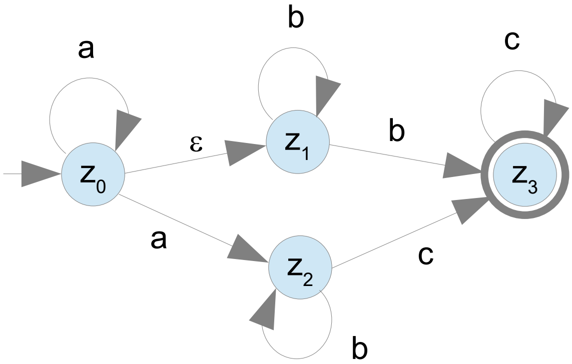
Dass man ε-NEAs in DEAs umwandeln kann, hat auch einen praktischen Nutzen: Man kann ε-NEAs verwenden, um zunächst "bequem" Sprachen zu definieren. Dann kann der Automat in einen DEA umgewandelt und somit auch effizient implementiert werden.

8.3.1 Teilmengenkonstruktion

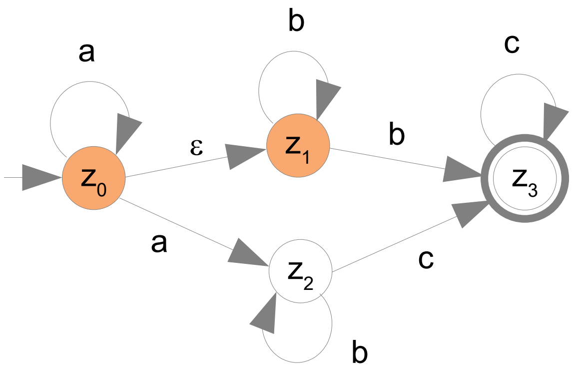
Wie Sie im Folgenden sehen werden, lässt sich jeder ε-NEA systematisch in einen äquivalenten DEA umwandeln. Die Grundidee der dazu verwendete sog. *Teilmengenkonstruktion* (oder auch *Potenzmengenkonstruktion*, da die Menge aller Teilmengen dabei verwendet wird) besteht darin, dass man Mengen von Zuständen betrachtet, in der sich der ε-NEA befinden könnte, nachdem er Teile eines Worts bereits verarbeitet hat.

Beispiel 8.11 - Erreichbare Zustände eines ε-NEA

Wir betrachten, wie sich folgender ε-NEA bei der Verarbeitung des Wort **ab** verhalten kann, d.h. welche Zustände dabei erreicht werden können.



- Ganz am Anfang, bevor das erste Zeichen **a** akzeptiert wird, kann der Automat noch im Startzustand z_0 sein oder schon spontan über ε-Übergang nach z_1 gewechselt haben. Die Menge möglicher, erreichbarer Zustände ist also $\{z_0, z_1\}$ (hier orange dargestellt):

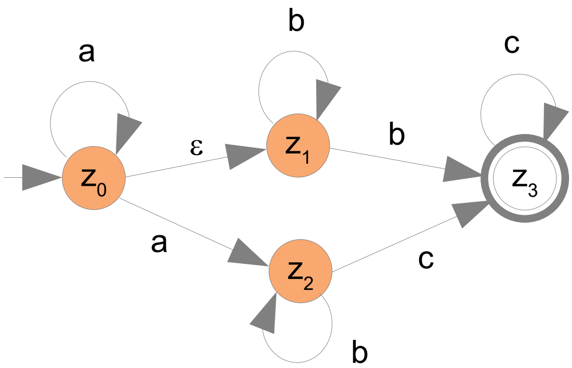


Eingabewort:

a	b

- Nach dem ersten Zeichen **a** wäre die Situation so:
  - Von z_0 könnte Automat **a** verarbeiten und entweder nach z_0 oder nach z_2 wechseln. Von z_0 könnte er dann auch gleich noch spontan nach z_1 wechseln.
  - Wäre der Automat in z_1 gewesen, könnte er mit **a** nichts anfangen.

  Nach dem ersten Zeichen **a** ist die Menge der Zustände, in der sich der Automat befinden könnte, also $\{z_0, z_1, z_2\}$.

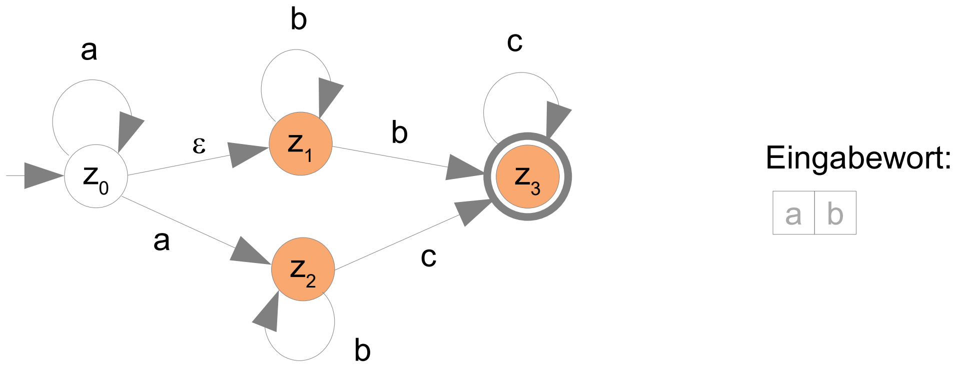


Eingabewort:

a	b

- Nach dem nächsten Zeichen **b** ergeben sich dann folgende Möglichkeiten:
  - Von z_0 aus geht es direkt mit **b** nicht weiter.
  - Von z_1 aus kann man mit **b** nach z_1 oder z_3 kommen.
  - von z_2 aus kann man mit **b** wieder nach z_2 kommen.

  Nach dem zweiten Zeichen **b** ist die Menge der Zustände, in der sich der Automat befinden könnte, die also mit **ab** erreichbar sind, die Menge $\{z_1, z_2, z_3\}$.



  Der Endzustand z_3 ist in der Menge der mit **ab** erreichbaren Zustände, d.h. **ab** wird von diesem ε-NEA akzeptiert.

Verfahren 8.12 - Teilmengenkonstruktion (Potenzmengen-Konstruktion)

Gegeben ist ein ε-NEA

$$A_N = (Z_N, \Sigma, \Delta, z_{N0}, E_N).$$

Der zugehörige äquivalente DEA

$$A_D = (Z_D, \Sigma, \delta_D, z_{D0}, E_D)$$

wird folgendermaßen definiert:

- **Zustandsmenge Z_D**: Menge aller Teilmengen von Z_N (= Potenzmenge von Z_N).
- **Zustandübergangsfunktion δ_D**: Ist $X = \{z_1, ..., z_k\}$ ein Zustand des DEA, dann ist $\delta_D(X,a)$ die Menge aller ε-NEA-Zustände, die ausgehend von einem der Zustände z_i durch einen direkten a-Übergang und dann beliebig viele ε-Übergänge im ε-NEA erreichbar sind.
- **Startzustand z_{D0}**: Menge aus dem Startzustand z_{N0} und allen Zuständen, die über ε-Übergänge von z_{N0} aus erreichbar sind.
- **Endzustände E_D**: Allen Zustandsmengen, die mindestens einen ε-NEA-Endzustand enthalten, sind Endzustände des DEA.

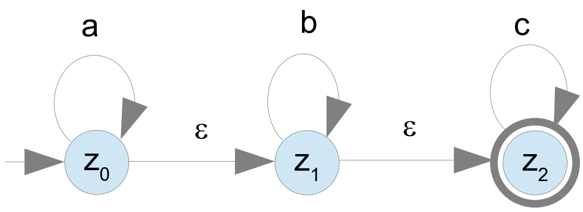
Anmerkung

- Es wird auf diese Weise pro Zustand für jedes Zeichen genau ein

Nachfolgezustand definiert, so dass ein deterministischer endlicher Automat (DEA) entsteht.

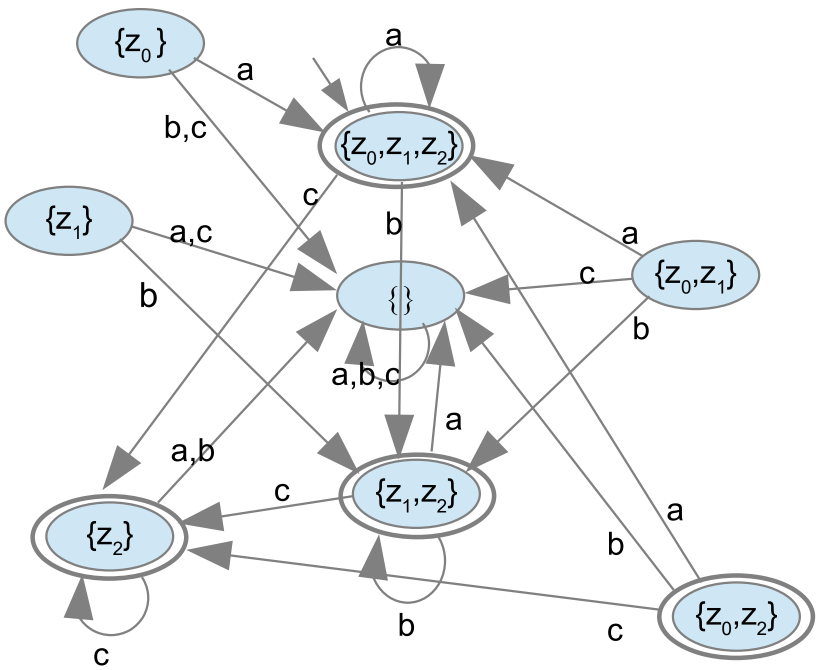
Beispiel 8.13 - Teilmengenkonstruktion

Folgender ε-NEA aus Beispiel 8.2 soll in einen DEA umgewandelt werden.



- Bei 3 Zuständen im ε-NEA gibt es $2^3 = 8$ Teilmengen als Zustände:
  $Z_D = \{ \{\}, \{z_0\}, \{z_1\}, \{z_2\}, \{z_0,z_1\}, \{z_0,z_2\}, \{z_1,z_2\}, \{z_0,z_1,z_2\} \}$
- Der Startzustand des DEA ist Zustand $\{z_0, z_1, z_2\}$, da z_0 der Startzustand des ε-NEA ist und z_1 und z_2 über ε-Übergänge erreichbar sind.
- Alle Teilmengen, die den Endzustand z_2 enthalten, sind Endzustände:
  $E_D = \{ \{z_2\}, \{z_0,z_2\}, \{z_1,z_2\}, \{z_0,z_1,z_2\} \}$
- Zur Bestimmung der Zustandsübergänge betrachten wir zunächst als Beispiel den Zustand $\{z_0,z_1,z_2\}$:
  - Mit Zeichen *a* und beliebig vielen nachfolgenden ε-Transitionen sind
    von z_0 aus z_0, z_1 und z_2 erreichbar,
    von z_1 aus nichts (*a* nicht im ersten Schritt verarbeitbar) und
    von z_2 (*a* nicht im ersten Schritt verarbeitbar) auch nichts.
    Somit ist von $\{z_0,z_1,z_2\}$ aus mit *a* wieder $\{z_0,z_1,z_2\}$ erreichbar.
  - Mit Zeichen *b* und beliebig vielen nachfolgenden ε-Transitionen sind
    von z_0 aus nichts (*b* nicht direkt möglich) und
    von z_1 aus z_1 und z_2
    von z_2 aus nichts erreichbar.
    Somit ist von $\{z_0,z_1,z_2\}$ aus mit *b* $\{z_1,z_2\}$ erreichbar.
  - Mit Zeichen *c* und beliebig vielen nachfolgenden ε-Transitionen sind
    von z_0 aus nichts und
    von z_1 aus nichts und
    von z_2 aus z_2 erreichbar.
    Somit ist von $\{z_0,z_1,z_2\}$ aus mit *c* $\{z_2\}$ erreichbar.

Insgesamt ergeben sich auf diese Weise folgende Zustandsübergänge:



Eigenschaft 8.14 - Äquivalenz von ε-NEA und konstruiertem DEA

Ist A_D der nach Verfahren 8.12 (Teilmengenkonstruktion) gebildete DEA zu einem ε-NEA A_N, dann gilt

$$L(A_D) = L(A_N),$$

d.h. beide Automaten akzeptieren die gleiche Sprache.

Auf den genauen Beweis dafür soll hier verzichtet werden. Es wäre folgendes nachzuweisen:

- Wir ein Wort *w* vom ε-NEA akzeptiert, d.h. kann man damit einen Endzustand erreichen, dann führt das Wort im konstruierten DEA auch zu einem Endzustand, d.h. dann gehört das Wort *w* auch zur Sprache des DEA.
- Wird ein Wort w vom DEA akzeptiert, d.h. erreicht man mit einem Wort *w* im konstruierten DEA einen Endzustand, dann gibt es im ε-NEA eine Möglichkeit, vom Startzustand mit *w* einen Endzustand zu erreichen. Somit gehört das Wort dann auch zur Sprache des ε-NEA.

Anmerkungen

- Hat der gegebene ε-NEA *n* Zustände, dann hat der durch die Teilmengenkonstruktion gebildete DEA 2^n Zustände, da eine Menge von *n*

Elementen 2^n Teilmengen hat. Bei einem ε-NEA mit 20 Zuständen hätte der DEA also 2^{20} = ca.1 Mio. Zustände! Meist (aber nicht immer) sind sehr viele Zustände des konstruierten DEA vom Startzustand aus nicht erreichbar und somit überflüssig.

- Der konstruierter DEA könnte durch anschließende Minimierung vereinfacht werden. Durch die vielen Zustände wäre aber die Minimierung äußerst aufwendig. Wir betrachten nachfolgend ein verbessertes Verfahren, das in den meisten Fällen weniger Zustände erzeugt und damit praktisch einsetzbar ist (z.B. in Scanner-Generatoren).

8.3.2 Verbesserte Teilmengenkonstruktion

Für praktische Anwendungen kann folgendes iterative Verfahren verwendet werden, um einen ε-NEA in einen äquivalenten DEA umzuwandeln, das meist deutlich weniger Zustände erzeugt als die 2^n Zustände nach Verfahren 8.12.

Man fängt dabei mit dem Startzustand an und nimmt neue Zustände "on demand" dazu, sobald sie gebraucht werden. Damit wird erreicht, dass nur erreichbare Teilmengen als Zustände verwendet werden. Durch eine Markierung wird gekennzeichnet, ob ein Zustand schon komplett bearbeitet ist und nicht mehr weiter betrachtet werden muss.

Verfahren 8.15 - Verbesserte Teilmengenkonstruktion

Gegeben ist ein ε-NEA $A_N = (Z_N, \Sigma, \Delta, z_{N0}, E_N)$. Der zugehörige äquivalente DEA $A_D = (Z_D, \Sigma, \delta_D, X_{D0}, E_D)$ wird folgendermaßen konstruiert:

1. **Initialisierung**: Bestimme Startzustand X_{D0}:

   X_{D0} besteht aus Startzustand z_{N0} und allen Zuständen, die über ε-Übergänge von z_{N0} aus direkt oder indirekt erreichbar sind.

   Setze $Z_D = \{ X_{D0} \}$, X_{D0} ist unmarkiert.

2. Verarbeitung der noch unmarkierten Zustände:
Wiederhole, solange noch ein unmarkierter Zustand in Z_D vorhanden ist:

(1) Wähle einen unmarkierten Zustand X aus Z_D aus.

(2) Für jedes Zeichen $a \in \Sigma$ berechne den Nachfolgezustand:

$X' = \{ z' \in Z_N \mid z'$ ist über erst einen a-Übergang und dann ggf. beliebig viele ε-Übergänge danach von einem Zustand $z \in X$ erreichbar$\}$

Nimm X' zu Z_D als unmarkierten Zustand dazu, falls noch nicht enthalten und setze

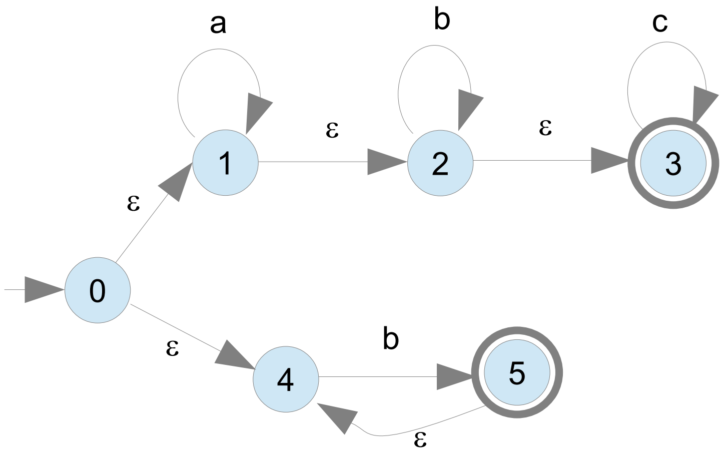
$\delta(X, a) = X'$

(3) Markiere Zustand *X* als fertig.

3. Endzustände: Alle Mengen aus Z_D, die mindestens einen Endzustand des ε-NEA enthalten, sind Endzustände des DEA.

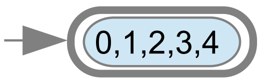
Beispiel 8.16 - Verbesserte Teilmengenkonstruktion

Folgender ε-NEA hat 6 Zustände. Die ursprüngliche Teilmengenkonstruktion nach Verfahren 8.13 würde $2^6 = 64$ Zustände ergeben.

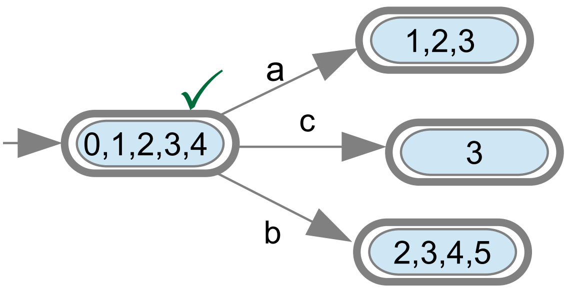


Verbesserte Teilmengenkonstruktion:

- Beginn mit Startzustand:

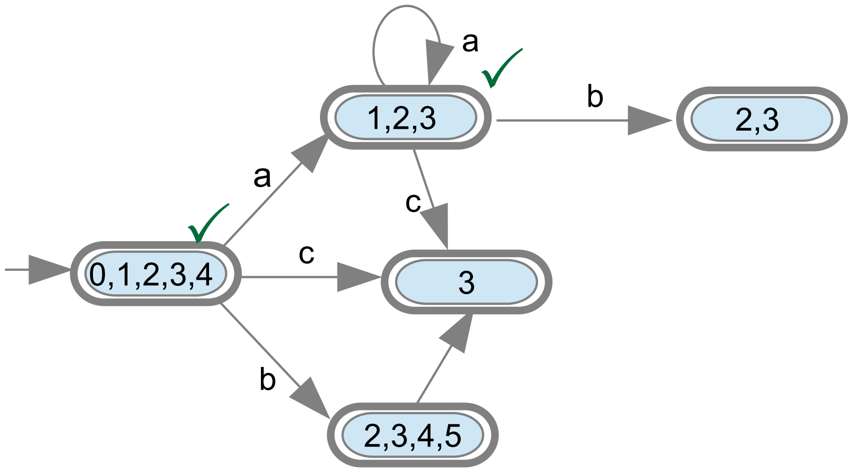


- Noch unmarkierter Startzustand wird zur Bearbeitung gewählt. Alle Nachfolger dafür bestimmen.



  Es sind drei weitere Zustände {1,2,3}, {2} und {2,3,4,5} dazu gekommen. Der bearbeitete Startzustand wird als fertig markiert.

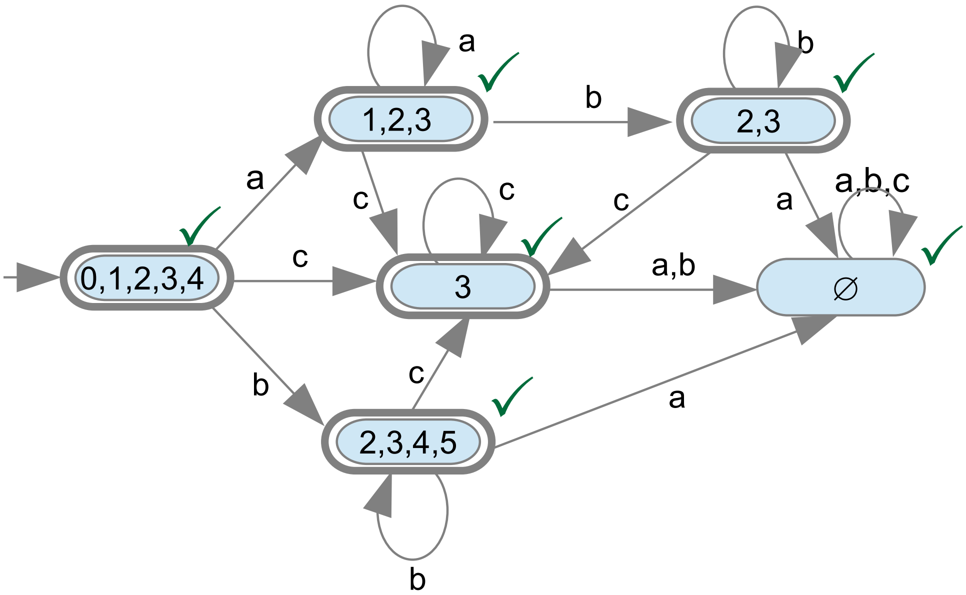
- Wähle einen noch unmarkierten Zustand zur Bearbeitung, z.B. {1,2,3}, bestimme Nachfolger für a, b und c und markiere Zustand als fertig. Es kommt dabei der Zustand {2,3} neu dazu.



Der bearbeitete Zustand {1,2,3} wird als fertig markiert.

- usw. …

Endergebnis:

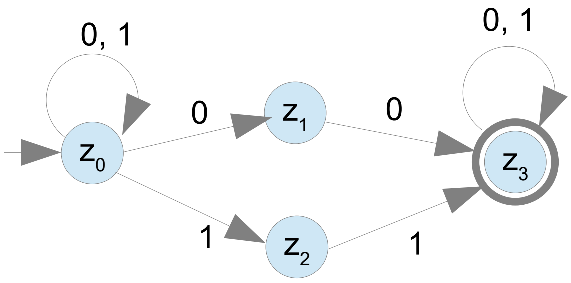


Anmerkungen

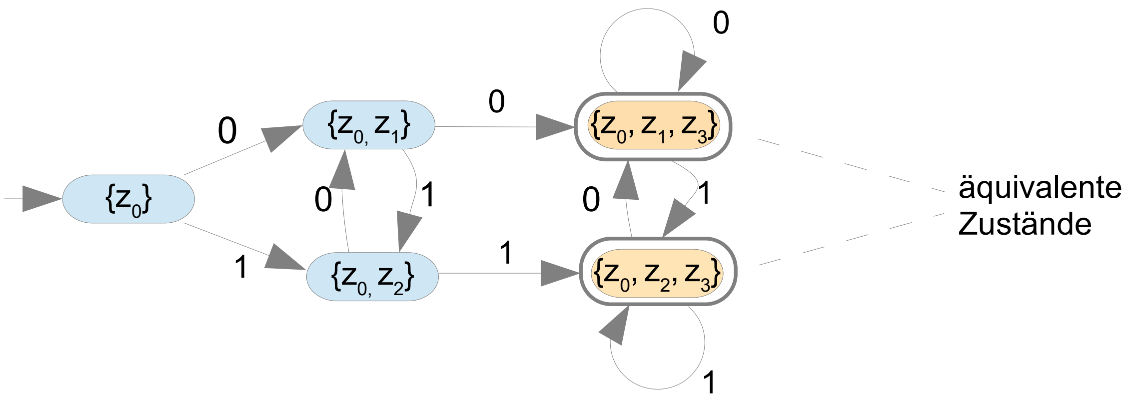
- Zu Beginn ist immer der Startzustand der einzige unmarkierte Zustand und wird somit als erstes bearbeitet.
- Auch die leere Zustandsmenge ∅ muss für jedes Zeichen einen Nachfolger haben (jeweils auch wieder die leere Menge).
- Die verbesserte Teilmengenkonstruktion liefert zwar meist einen Automaten mit deutlich weniger als 2^n Zuständen, aber nicht unbedingt einen Minimal-automaten!

Beispiel 8.17 - Verbesserte Teilmengenkonstruktion

Folgendes ist ein ε-NEA für Binärfolgen, die 00 oder 11 enthalten:

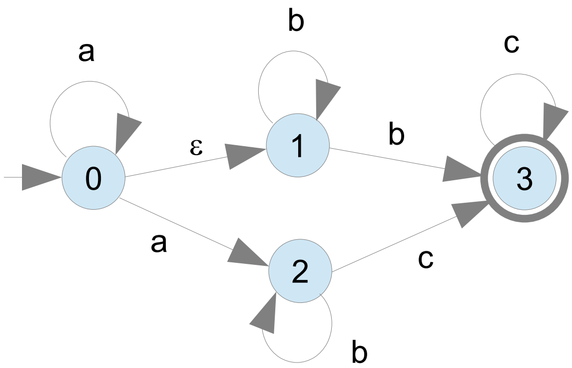


Der mit der verbesserten Teilmengenkonstruktion berechnete DEA enthält zwei äquivalente Zustände und ist somit nicht minimal.



Aufgabe 8.18 - Verbesserte Teilmengenkonstruktion

Wandeln Sie mit der verbesserten Teilmengenkonstruktion folgenden ε-NEA in einen DEA um.



Zusammenfassung zu Lektion 13

Diese Fragen sollten Sie nun beantworten können

- Wie sind nichtdeterministische endliche Automaten mit ε-Übergängen (ε-NEA) definiert?
- Wann wir ein Wort von einem ε-NEA akzeptiert?
- Wie können ε-NEA eingesetzt werden, um Sprachen zu beschreiben?
- Welcher Zusammenhang besteht zwischen ε-NEA und DEA?
- Wie ist die Teilmengenkonstruktion definiert, mit der ein ε-NEA in einen äquivalenten DEA umgewandelt werden kann?
- Wie wird mit der verbesserten Teilmengenkonstruktion zu einem ε-NEA ein äquivalenter DEA gebildet? Welchen Vorteil hat dieses Verfahren?