Institut für Informatik
Lehrstuhl V 02.06.00
Matr.Nr.: 702502 und 702336

Theoretische Informatik

Skript zur Vorlesung vom 12.05.2000

Zu Thema 4.

Theoretische Informatik

Skript zur Vorlesung vom 12.05.2000

1. Definition des deterministischen endlichen Akzeptors:

Allgemeine Definition:

Ein endlicher Akzeptor ist ein erkennender Akzeptor für eine Sprache L über einem Alphabet. Für ein Wort w aus dem Alphabet stellt er fest, ob das Wort in der Sprache L ist.

Solch ein Akzeptor hat einen Lesekopf, um eine Eingabe zu lesen und einen internen Zustand. Er darf das Eingabewort nur einmal von links nach rechts lesen. Die Anzahl der Zustände muß endlich sein. Einige der Zustände sind terminal. Wenn er sich nach dem Lesen des letzten Buchstaben von einem Wort w in einem terminalen Zustand befindet, akzeptiert er das Wort, d.h. er entscheidet, dass das Wort in L liegt, ansonsten entscheidet er, dass dieses Wort nicht zu L gehört. Bei einem Eingabewort stopt der Akzeptor auf jeden Fall nach |w| (lese)-Schritten und sagt dann „ja“ (terminalerZustand) oder „nein“ (nichtterminaler Zustand). Ein gutes Beispiel für einen endlichen Akzeptor im täglichen Leben ist ein Getränkeautomat. In seinem Zustand muß er sich verschiedene Dinge merken, z.B. wieviel Geld die Kundin oder der Kunde schon eingeworfen hat, welche Tasten gedrückt wurden und wieviel Flaschen welcher Getränkesorte noch im Automat vorhanden sind. Er hat also einen endlich großen Speicher. Die Größe wird zum Zeitpunkt der Konstruktion festgelegt.

DA ist die Klasse aller Sprachen, die von einem endlich determiniertem Akzeptor A akzeptiert wird. Formal wird der Akzeptor durch ein 5-tupel beschrieben.

A=(X,S,δ,s0, F)

wobei

X,S endliche Mengen sind (X- Alphabet, S- Zustandsmenge)

δ ist die Zustandsüberführungsfunktion mit δ:SxX$\rightarrow$S

S0 $\in$ S

F ist die Menge der terminalen Zustände, F$\in$S

Um mehrere Übergangsschritte auf einmal betrachten zu können, erweitern wir δ auf δ*. Eine Mehrfachanwendung der Übergangsfunktion δ*:SxX*$\rightarrow$S ist induktiv über X* definiert wie folgt:

für alle a$\in$X und alle w$\in$X*, alle s,s0$\in$S

(δ*(s,ε)=s;

δ*(s0,wa)=δ(δ*(s,w),a)

Wenn klar ist was gemeint ist, wird δ* auch einfach als δ geschrieben.

DA ist die Klasse aller Sprachen, die von einem endlich determiniertem Akzeptor A akzeptiert wird.

- **Charakterisierung**
  - $x\in$ DA => x erfüllt eine Eigenschaft E
  - $x\in$ DA <= x erfüllt eine Eigenschaft E‘
  - Daraus folgt die Äquivalenz: $x\in$ DA ⇔ x erfüllt eine Eigenschaft E
- **Abschlußeigenschaften**: **Vereinigung**

  Durchschnitt (siehe Hausaufgabe)

  Differenz

  Komplement

1.1. Satz D:

DA ist abgeschlossen gegen Vereinigung, Durchschnitt, Komplement und Differenz von Mengen: R1,R2 $\in$ DA => R1$\cap$R2$\in$DA, R1$\cup$R2$\in$DA, R1\R2$\in$DA, $\overline{R1}\in$ DA
Beweis:

1.1.1. Komplement:

Komplement von A ist $\overline{A}$; $\overline{A}$=(X1,S1,δ1,s01,S1\F) $\overline{A}$ akzeptiert die Wörter, die A nicht akzeptiert.
zu zeigen ist

$$L(\overline{A}) \subseteq \overline{R1} = \overline{L(A)} \wedge L(\overline{A}) \supseteq \overline{R1} = \overline{L(A)}$$

Hierbei kann man die beiden Beweisrichtungen zusammenfassen.

w$\in$ $L(\overline{A}) \Leftrightarrow \delta(s_{01}, w) \in$ S\F $\Leftrightarrow \delta(s_{01}, w) \in$ F $\Leftrightarrow$ w$\in$ X*\R1 $\Leftrightarrow$ w$\in$ $\overline{R}$

Also ist L(A)=X*\R$\in$ DA. q.e.d.

1.1.2. Vereinigung:

Sei R1,R2 aus DA und es existieren 2 Akzeptoren A1,A2 mit L(A1)=R1 und L(A2) = R2.

mit A1=(X1,S1,δ1,s01,F1) und A2=(X2,S2,δ2,s02,F2)

Behauptung: DA ist abgeschlossen gegenüber der „Vereinigung“ $\cup$.

Für alle R1,R2 aus DA ist R1$\cup$R2 $\in$ DA.

Definition der Vereinigung: M$\cup$N := {y|y$\in$M **oder** y$\in$N} wobei M,N beliebige Mengen sind.

Beweis:

Sei R1,R2 aus DA und es existieren 2 Akzeptoren A1,A2 mit L(A1)=R1 und L(A2) = R2. mit A1=(X1,S1,δ1,s01,F1) und A2=(X2,S2,δ2,s02,F2)

Ich konstruiere einen Akzeptor:

Sei A=(X,S,δ,(s01,s02),F) ein endlicher determinierter Akzeptor mit

X= X1 ∪ X2 und S=(S1∪{q1}) x (S2∪{q2}) wobei q1 bzw. q2 aus S1 bzw. S2 ist.

In dieser Akzeptor werden die beiden Akzeptoren A1 und A2 „quasi parallel" abgearbeitet.

In den Zuständen q1und q2 werden die Sonderfälle abgefangen:

(d.h. Die Sonderfälle sind die Fälle, in denen der Automat in einen nichtdefinierten Zustand übergehen würde.)

δ1(s,x) =q1 für x nicht Element X1 **oder** δ2(s',x) = q2 für x nicht Element X2 ; s,s'∈S

sowie δ1(q1,x) = q1; δ2(q2,x)=q2 für alle x aus X

δ((s,s'),x) = (δ1(s,x),δ2(s',x)) für alle x∈X, s∈S1 und s'∈S2.

F={(s,s') ∈S | s∈F1 **oder** s' ∈F2} (nach Definition der Vereinigung.)

Wenn einer der beiden Akzeptoren in den Zustand q1 bzw. q2 übergehen, ist (s,s') ∉ F , d.h. der Akzeptor A kann nicht mehr in einen terminalen Zustand übergehen. Es wird ein paralleles Arbeiten der 2 Akzeptoren durch Übergang ins Kreuzprodukt der Zustände geschaffen. Die Vereinigung der Sprachen R1,R2 wird akzeptiert, wenn einer der beiden Akzeptoren die jeweilige Sprache akzeptiert.

Jetzt ist zu zeigen: L(A) = R1∪R2 = L(A1) ∪ L(A2)

Dazu ist noch zu beweisen:

δ((s,s'),w)= (δ1(s,w),δ2(s',w)) für alle w∈ X*; s∈S1 und s'∈S2

Dies werde ich über vollständige Induktion über |w| zu zeigen haben:

Induktionsvoraussetzung: |w| = 0; d.h. w=ε

δ((s,s'),ε) = (δ1(s, ε),δ2(s',ε))

(s,s') = (s,s') - ist eine wahre Aussage.

Induktionsanfang: |w|=n mit n∈N

δ((s,s'),w)= (δ1(s,w),δ2(s',w))

$w = w_1 \circ w_2 \circ w_3 \circ \ldots \circ w_n$

Induktionsschritt: w'∈X |w'| = n+1 n∈N

$w' = w_1 \circ w_2 \circ w_3 \circ \ldots \circ w_n \circ w_{n+1} = w \circ x$

mit w_{n+1}=x; $w_{n+1} \in X^*$; $x \in X^*$;

δ((s,s'),w')= (δ1(s,w'),δ2(s',w')) kann man auch schreiben als

δ((s,s'),wx)=(δ1(s,wx),δ2(s',wx))

δ((s,δ((s,s'),w),x)= (δ1(δ1(s,w),x),δ2(δ2(s',w),x))

ich setzte (s,s')=($s_{i,}$ s'_i) mit $s_{i,} \in$ S1;$s'_i \in$S2;

Man wende jetzt den Induktionsanfang an.
Nach Eingabe von w aus X* (ein ganzes Wort) ist man im Zustand (s_i, s'_i),
mit i=|w|, wenn (s_i, s'_i) der Zustand vor Eingabe des Wortes w ist.

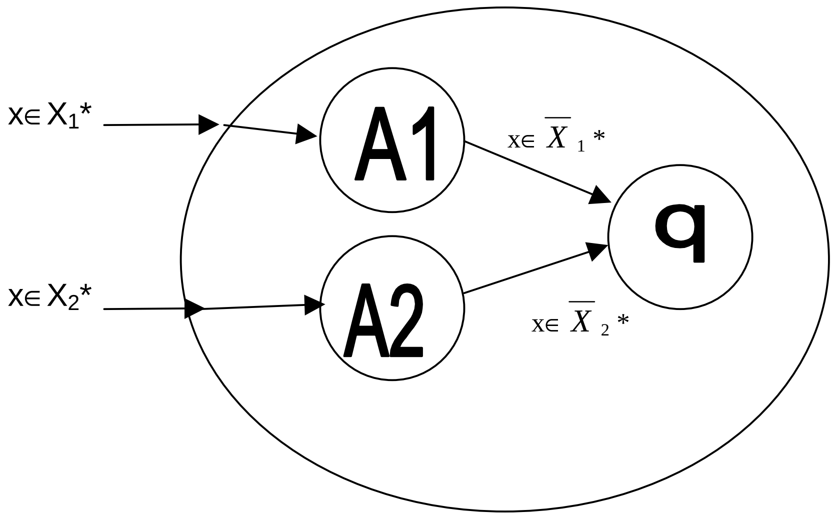
$\delta((s_i,s'_i),x)=(\delta 1(s_i,x),\delta 2(s'_i,x))$

nach der Anwendung von einem weiterem Buchstaben ist man ein Zustand weiter.
$\delta((s_{i+1},s'_{i+1}),\varepsilon)= (\delta 1(s_{i+1}, \varepsilon),\delta 2(s'_{i+1}, \varepsilon))$
Das ergibt nach Induktionsvoraussetzung
$(s_{i+1},s'_{i+1})= (s_{i+1},s'_{i+1})$ *Das ist eine wahre Aussage*.
Hier wurde bewiesen, daß
$\delta((s,s'),w)= (\delta 1(s,w),\delta 2(s',w))$. q.e.d.

Wenn $w\in L(A)$ ist, dann ist $\delta((s01,s02),w)\in F$, d.h. bei Eingabe des Wortes w gelangt der Akzeptor in einen Endzustand.
Genau dann ist $\delta 1(s01,w)\in F1$ **oder** $\delta 2(s02,w)\in F2$ (nach obigen Beweis durch Vollständige Induktion)
$w\in R1=L(A1)$ **oder** $w\in R2=L(A2)$ ⇔ **$w\in R1\cup R2$**
⇔ $R1,R2 \in DA$ Abgeschlossen gegen Vereinigung.
Also ist **$L(A)=R1\cup R2 \in DA$**

q.e.d.

Diese Skizze soll das Zusammenspiel zwischen den Akzeptoren A1 und A2 demonstrieren.



1.1.3. Die Differenz:

Die Differenz besteht nach Definition aus dem Durchschnitt und dem Komplement.
R1\R2 = R1 $\cap$ (X2*\R2)
(X2*\R2) = $\overline{R2}$ Das Komplement haben wir schon bewiesen.
Den Durchschnitt haben wir auch schon bewiesen.
Also führt die Differenz auch nicht aus der Menge raus.

2. Nicht deterministische endliche Akzeptoren

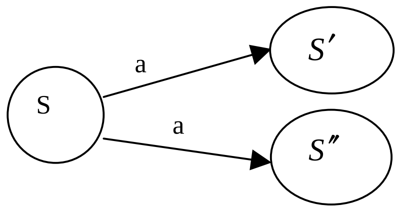
„Nichtdeterminismus“: $\approx$ “Unbestimmtheit, Willkürlichkeit, Spontanität“

Nichtdeterministische Akzeptoren:

Merkmal:

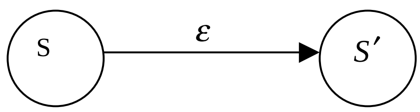
willkürlicher Zustandsübergang:

Das heißt: mit dem Lesen eines Zeichens besteht die Möglichkeit mit ein und dem selben Zeichen in zwei verschiedene Folgezustände überzugehen.



a $\in$ X*

Es besteht die Möglichkeit ohne lesen eines Zeichens von einen Zustand in einen anderen Zustand mit Hilfe eines(ε - Übergang) überzugehen.



2.1. Def. F :

A=(X,S,δ,S $_0$,F) sei ein nichtdeterministischer endlicher Akzeptor

Wenn:

X,S sind endliche nichtleere Mengen

δ : S$\times$X $\rightarrow$ 2^S (Überführungsrelation)

(2^S $\rightarrow$ Menge aller Teilmengen von S)

$S_0 \subseteq$ S Menge der Anfangszustände

F $\subseteq$S Menge der Endzustände

Falls S_0 nur ein Zustand s_0 enthält schreiben wir oft A=(X,S,δ,s_0 ,F)

A heißt vollständig, wenn δ : S$\times$X $\rightarrow$ 2^S \{$\emptyset$}, d.h. wenn es zu jedem Zustand s$\in$ S' und jeder Eingabe X mindestens einen Folgezustand $s' \in$ δ (S, x) gibt.

Ein nicht deterministischer endlicher Automat heißt determiniert, wenn er nie eine Auswahlmöglichkeit hat, d.h. rein formal können nicht deterministische Automaten also auch determiniert sein.

2.2. Def. G:

Fortsetzung von δ auf Wörter

(d.h. Erweiterung von δ von Buchstaben auf Wörter dazu muss man die Relation δ^* : S $\times$X* $\rightarrow$$2^S$ definieren)

Sei A=(X,S, δ,s_0 ,F) ein nichtdeterministischer endlicher Akzeptor.

Dann setze man δ zu einer Relation δ^* : S $\times$X* $\rightarrow$$2^S$

Somit folgt : δ^* (s,ε) ={s} $\forall$ s$\in$ S und ε ist das leere Wort

$$\delta^*(s,xw)= \bigcup_{t\in\delta(s,x)} \delta^*(t,w)\ ,\ \forall s\in S,\ x\in X,\ w\in X^*$$

δ^* wird auch als δ geschrieben, da es im Kontext ersichtlich ist.

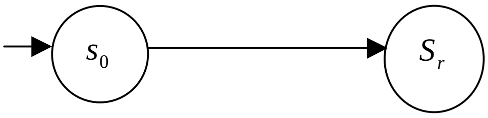
Erkannte Sprache von A wird definiert über alle Wege von einen Anfangszustand zu einem Endzustand.

2.3. Def. H

Sei A=(X,S, δ ,s_0 ,F) ein nichtdeterministischer endlicher Akzeptor. Die Menge L(A)= { w$\in$ X^* | Es gibt ein $s_0 \in S_0$ mit δ (s_0 ,w)$\cap$F$\neq\emptyset$ } ist die von A akzeptierte Sprache.

NA ist definiert als die Klasse aller Sprachen, die von nicht deterministischen endlichen Akzeptoren akzeptiert werden,

NVA diejenige Klasse, die von nichtdeterministischen endlichen vollständigen Akzeptoren erkannt wird.



r$\in$ N

Ein nicht deterministischer endlicher Automat A akzeptiert ein Wort w, wenn es zu w mindestens einen Weg mit der „Beschriftung „ von w durch A gibt ,d.h. es gibt eine Kantenfolge welche mit den Buchstaben des Wortes w beschriftet ist, diese Kantenfolge bzw. der Weg muss in einen terminalen Zustand enden. Neben diesen einen erfolgreichen Weg darf es noch beliebig viele Sackgassen geben. Man muss nur sicherstellen, dass der Automat auf keinen Fall nach dem Lesen eines Wortes w' in einen terminalen Zustand ist, wenn w' nicht zur Sprache L(A) gehört.

Man kann nicht determinierende Automaten auf zwei Weisen betrachten: Entweder man sagt, der Automat kann raten, welchen von mehreren möglichen Folgezuständen er wählt. Eine andere Sichtweise ist, dass er alle Wege mit „Beschriftung“ w parallel durchläuft. Diese Sichtweise kann man einnehmen, da ja ein Wort w akzeptiert wird, wenn auch nur einer der möglichen Zustände nach Lesen von w ein terminaler Zustand ist. Wenn es in einem nichtdeterministischen Akezeptor keinen möglichen Folgezustand gibt, so entspricht das einem „absorbierenden Zustand“ in einem deterministischen endlichen Akzeptor, von dem aus kein Weg mehr zu einem terminalen Zustand führt: Das Eingabewort wird nicht akzeptiert.

Beispiel: Texterkennung von oben

Sei X beliebiges Alphabet und $v = x_1 \ldots x_r \in X^*$ mit $x_i \in X$ für i=1,...,r $r \in N$, $i \in N$

$L_v = \{wv \mid w \in X^*\}$

Definiere

Sei $A_v = (x, \{s_0, .., s_r, q\}, \delta, \{s_0\}, \{s_r\})$ ein nichtdeterministischer endlicher Akzeptor.

Übergangsfunktion:

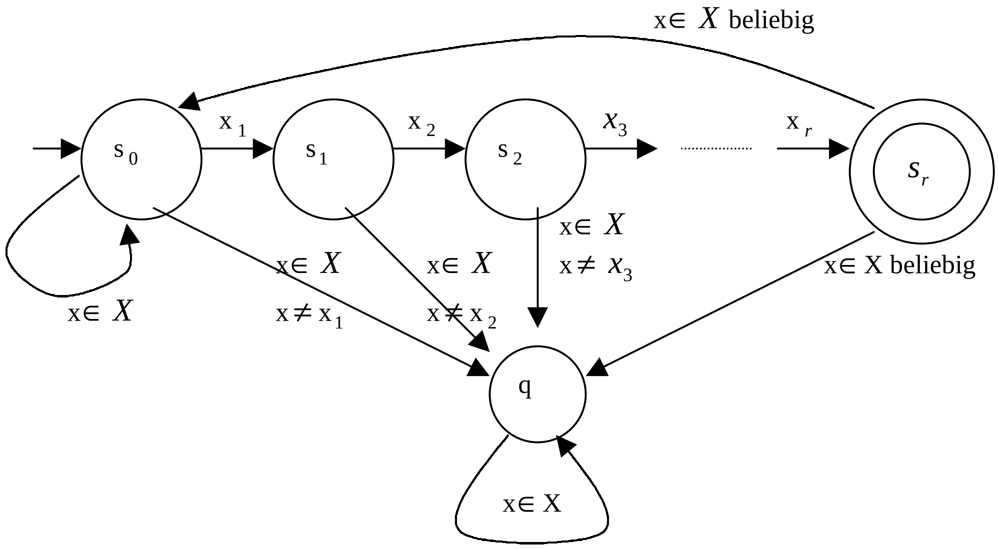
$\delta(s_0, x) = \{s_0, q\}$ für $x \neq x_1$

$\delta(s_0, x_1) = \{s_0, s_1\}$

$\delta(s_i, x) = \{q\}$ für $x \neq x_{i+1}$ und i=1,....r-1

$\delta(s_r, x) = \{s_0, q\}$ für alle $x \in X$

$\delta(s_i, x_{i+1}) = \{s_{i+1}\}$, i= 1, …, r-1



Erläuterung:

Wenn der Endzustand erreicht wird , wurde ein Wort erkannt, ansonsten geht der Automat in den absorbierenden Zustand q. in diesem Fall entspricht das Wort nicht dem gesuchten Wort.

Es gibt einen Weg von s_0 nach s_r

Offenbar: $L(A) = L_v$

2.4. Satz I : DA = NVA = NA

Beweis:
Offenbar DA $\subseteq$NVA$\subseteq$NA, weil ein deterministischer endlicher Akzeptor auch ein nichtdeterministischer vollständiger Akzeptor und der wiederum ein allgemeiner nichtdeterministischer endlicher Akzeptor ist.

„NA $\subseteq$NVA“ :

Sei A=(X,S, δ ,s_0 ,F) ein nichtdeterministischer endlicher Akzeptor.
Bilde A'=(X,S$\cup${q},δ',s_0,F) durch Ergänzung eines neuen absorbierenden Zustandes

q$\notin$ S und $\delta'(s,x)=\begin{cases} \delta(s,x), \mathit{falls} \delta(s,x) \neq \emptyset \\ \{q\}, \mathit{falls} \delta(s,x) = \emptyset \quad oder \quad falls \quad s = q \end{cases}$ x$\in$ X, s$\in$ S

L(A)=L(A') $\Rightarrow$NVA$\supseteq$NA

“NVA$\subseteq$DA“:
(Beweis nächstes Vorlesungsscript)