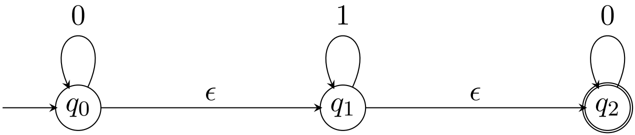
5.2 NFA's mit ϵ-Übergängen

Definition 65

Ein (nichtdeterministischer) endlicher Automat A mit ϵ-Übergängen ist ein 5-Tupel analog zur Definition des NFA mit

$$\delta : Q \times (\Sigma \uplus \{\epsilon\}) \rightarrow \mathcal{P}(Q) \ .$$

Ein ϵ-Übergang wird ausgeführt, ohne dass ein Eingabezeichen gelesen wird. Wir setzen o.B.d.A. voraus, dass A nur einen Anfangszustand hat.



Definiere für alle $a \in \Sigma$

$$\bar{\delta}(q, a) := \hat{\delta}(q, \epsilon^* a \epsilon^*) \; .$$

Falls A das leere Wort ϵ mittels ϵ-Übergängen akzeptiert, also $F \cap \hat{\delta}(q_0, \epsilon^*) \neq \emptyset$, dann setze zusätzlich

$$F := F \cup \{q_0\} \; .$$

Satz 66

$$w \in L(A) \Leftrightarrow \hat{\hat{\delta}}(S, w) \cap \; F \neq \emptyset \; .$$

Beweis:
Hausaufgabe! □

5.3 Entfernen von ϵ-Übergängen

Satz 67

Zu jedem nichtdeterministischen endlichen Automaten A mit ϵ-Übergängen gibt es einen nichtdeterministischen endlichen Automaten A' ohne ϵ-Übergänge, so dass gilt:

$$L(A) = L(A')$$

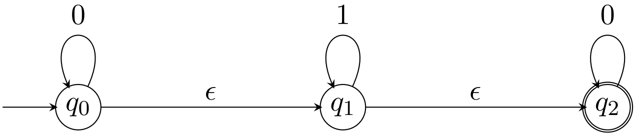
Beweis:

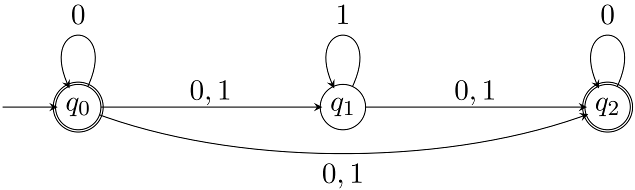
Ersetze δ durch $\bar{\delta}$ und F durch F' mit

$$F' = \begin{cases} F & \epsilon \notin L(A) \\ F \cup \{q_0\} & \epsilon \in L(A) \end{cases}$$

$\square$

Beispiel 68





5.4 Endliche Automaten und reguläre Sprachen

Satz 69

Ist $G = (V, \Sigma, P, S)$ eine rechtslineare (also reguläre) Grammatik (o.B.d.A. sind die rechten Seiten aller Produktionen aus $\Sigma \cup \{\epsilon\} \cup (\Sigma \cup \{\epsilon\})V$), so ist $N = (V \cup \{X\}, \Sigma, \delta, \{S\}, F)$, mit

$$F := \begin{cases} \{S, X\}, & \text{falls } S \to \epsilon \in P \\ \{X\}, & \text{sonst} \end{cases}$$

und, für alle $A, B \in V$, $a \in \Sigma \cup \{\epsilon\}$,

$$\begin{array}{lcll} B \in \delta(A, a) & \Longleftrightarrow & A \to aB & \text{und} \\ X \in \delta(A, a) & \Longleftrightarrow & A \to a & \end{array}$$

ein nichtdeterministischer endlicher Automat, der genau $L(G)$ akzeptiert.

Beweis:

Aus der Konstruktion folgt, dass N ein NFA ist (i.A. mit ϵ-Übergängen, die sich bei Produktionen der Form $A \to B$ ergeben).

Durch eine einfache Induktion über n zeigt man, dass eine Satzform

$$a_1 a_2 \cdots a_{n-1} A \text{ bzw. } a_1 a_2 \cdots a_n$$

in G genau dann ableitbar ist, wenn für die erweiterte Übergangsfunktion $\hat{\delta}$ des zu N äquivalenten NFA *ohne* ϵ-Übergänge gilt:

$$A \in \hat{\delta}(S, a_1 a_2 \cdots a_{n-1})$$

bzw.

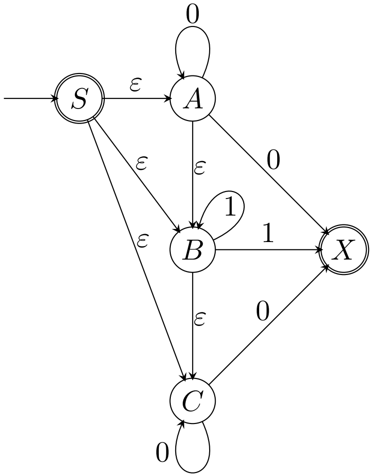
$$X \in \hat{\delta}(S, a_1 a_2 \cdots a_n)$$

(bzw., für $n = 0$, $F \cap \hat{\delta}(S, \epsilon) \neq \emptyset$). $\square$

Beispiel 70

Wir betrachten die (reguläre) Grammatik G mit den Produktionen

$$
\begin{array}{lcl}
S & \rightarrow & \varepsilon \\
S & \rightarrow & A \mid B \mid C \\
A & \rightarrow & 0A \mid 0 \mid B \\
B & \rightarrow & 1B \mid 1 \mid C \\
C & \rightarrow & 0C \mid 0
\end{array}
$$



Zusammenfassend ergibt sich:

Satz 71

Die Klasse der regulären Sprachen (Chomsky-3-Sprachen) ist identisch mit der Klasse der Sprachen, die

- *von DFA's akzeptiert/erkannt werden,*
- *von NFA's akzeptiert werden,*
- *von NFA's mit ϵ-Übergängen akzeptiert werden.*

Beweis:

Wie soeben gezeigt. $\square$

5.5 Reguläre Ausdrücke

Reguläre Ausdrücke sollen eine kompakte Notation für spezielle Sprachen sein, wobei endliche Ausdrücke hier auch unendliche Mengen beschreiben können.

Definition 72

Reguläre Ausdrücke sind induktiv definiert durch:

1. $\emptyset$ ist ein regulärer Ausdruck.
2. ϵ ist ein regulärer Ausdruck.
3. Für jedes $a \in \Sigma$ ist a ist ein regulärer Ausdruck.
4. Wenn α und β reguläre Ausdrücke sind, dann sind auch (α), $\alpha\beta$, $(\alpha|\beta)$ (hierfür wird oft auch $(\alpha + \beta)$ geschrieben) und $(\alpha)^*$ reguläre Ausdrücke.
5. Nichts sonst ist ein regulärer Ausdruck.

Bemerkung: Ist α atomar, so schreiben wir statt $(\alpha)^*$ oft auch nur α^*.

Zu einem regulären Ausdruck γ ist die zugehörige Sprache $L(\gamma)$ induktiv definiert durch:

Definition 73

1. Falls $\gamma = \emptyset$, so gilt $L(\gamma) = \emptyset$.
2. Falls $\gamma = \epsilon$, so gilt $L(\gamma) = \{\epsilon\}$.
3. Falls $\gamma = a$, so gilt $L(\gamma) = \{a\}$.
4. Falls $\gamma = (\alpha)$, so gilt $L(\gamma) = L(\alpha)$.
5. Falls $\gamma = \alpha\beta$, so gilt
$$L(\gamma) = L(\alpha)L(\beta) = \{uv;\ u \in L(\alpha), v \in L(\beta)\}\ .$$
6. Falls $\gamma = (\alpha \mid \beta)$, so gilt
$$L(\gamma) = L(\alpha) \cup L(\beta) = \{u;\ u \in L(\alpha) \vee u \in L(\beta)\}\ .$$
7. Falls $\gamma = (\alpha)^*$, so gilt
$$L(\gamma) = (L(\alpha))^* = \{u_1 u_2 \dots u_n;\ n \in \mathbb{N}_0, u_1, \dots, u_n \in L(\alpha)\}\ .$$

Beispiel 74

Sei das zugrunde liegende Alphabet $\Sigma = \{0, 1\}$.

- alle Wörter, die gleich 0 sind oder mit 00 enden:

$$(0 \mid (0 \mid 1)^* 00)$$

- alle Wörter, die 0110 enthalten:

$$(0|1)^* 0110 (0|1)^*$$

- alle Wörter, die eine gerade Anzahl von 1'en enthalten:

$$(0^* 1 0^* 1)^* 0^*$$

- alle Wörter, die die Binärdarstellung einer durch 3 teilbaren Zahl darstellen, also

$$0, 11, 110, 1001, 1100, 1111, 10010, \ldots$$

  Hausaufgabe!