Automatentheorie und Formale Sprachen

gehalten von Univ.-Prof. Dr. rer. nat. Peter Rossmanith
im Sommersemester 2005 an der RWTH Aachen

eine studentische Mitschrift von
Thomas Siegbert
thomas.siegbert@t-online.de

Letzte Änderung: 24. April 2005

Inhaltsverzeichnis

1 Formale Sprachen

1.1 Alphabete, Wörter, Sprachen

12.04.2005

Ein **Alphabet** ist eine endliche Menge von **Symbolen**.

Beispiel: {a,...,z} , {0,1} , ASCII

Was ist ein Wort, was ist eine Sprache ?

Informelle Antwort:

1. Ein Wort ist eine Aneinanderkettung von Symbolen.
2. Eine Sprache ist eine Menge von Wörtern.

Beispiel: 01, 101001, ϵ sind Wörter über dem Alphabet {0,1}.

Definition:

Eine Halbgruppe (H, ○) besteht aus einer Menge H und einen
assoziierten Operator ○: H x H → H.
Ein **Monoid** ist eine Halbgruppe mit einem neutralem Element.
Es sei (M, ○) ein Monoid und $E \subseteq M.$ E ist Erzeugendensystem von (M, ○), falls jedes $m \in M$ als $m = e_1 \circ e_2 \circ ... \circ e_n$ mit $e_i \in E$ dargestellt werden kann.

Beispiel: $(\mathbb{Z}, +)$ ist ein Monoid. {-1, 1} ist ein Erzeugendensystem davon.

Frage: Ist {-16, 13, 12} ein Erzeugendensystem für $(\mathbb{Z}, +)$? (//ja, es ist ein Erz.-System)

Ein Erzeugendensystem E für ein Monoid (M, ○) ist **frei**, falls jedes $m \in M$ auf nur eine Art als $m = e_1 \circ ... \circ e_n$ mit $e_i \in E$ dargestellt werden kann.

Beispiel:

- $(\mathbb{Z}, +)$ ist von {-1, 1} **nicht** frei erzeugt, denn: 2 = 1 + 1 = 1 + 1 + (-1) + 1
- $(\mathbb{N}_{0,} +)$ ist von {1} frei erzeugt,

$$0 = \sum_{i \in \emptyset} 1 \left(= \sum_{i=0}^{-1} a_n = 0 \right) \quad .$$

Satz:

Es sei Σ ein Alphabet. Dann ist das von Σ frei erzeugte Monoid bis auf Isomorphie eindeutig.
Zwei Monoide $(M_1, \bullet)$ und $(M_2, \circ)$ sind isomorph, falls es eine Abbildung
h: $M_1 \rightarrow M_2$ gibt mit

1. h ist bijektiv
2. h ist ein Homomorphismus, d.h. h(u●v) = h(u) ○ h(v) für alle $u, v \in M_1$.

Beweis:

Es seien $(M_1, \bullet)$ und $(M_2, \circ)$ zwei von Σ frei erzeugte Monoide. Wir definieren eine Abbildung h: $M_1 \rightarrow M_2$ auf folgende Weise:

$$h(w) = h(w_1 \bullet \ldots \bullet w_n) = w_1 \circ \ldots \circ w_n \quad \text{wobei} \quad w_i \in \Sigma \quad \text{ist.}$$

Behauptung: h ist ein Isomorphismus.

Wir zeigen zuerst, dass h bijektiv ist. Definiere g: $M_2 \rightarrow M_1$ vermöge $g(v) = v_1 \bullet \ldots \bullet v_n$ wobei $v = v_1 \circ \ldots \circ v_n$ mit $v_i \in \Sigma$.

$$g(h(w)) = g(h(w_1 \bullet \ldots \bullet w_n)) = g(w_1 \circ \ldots \circ w_n) = w_1 \bullet \ldots \bullet w_n = w$$

Nun ist außerdem

$$h(u \bullet v) = h(u_1 \bullet \ldots \bullet u_n \bullet v_1 \bullet \ldots \bullet v_m) = u_1 \circ \ldots \circ u_n \circ v_1 \circ \ldots \circ v_m$$

$$= (u_1 \circ \ldots \circ u_n) \circ (v_1 \circ \ldots \circ v_m) = h(u) \circ h(v) \quad .$$

□

Definition:

Es sei Σ ein Alphabet. Dann ist $(\Sigma^*, \cdot)$ das von Σ frei erzeugte Monoid.
Die Elemente von Σ^* nennen wir **Wörter** (über Σ).
Falls $L \subseteq \Sigma^*$, dann nennen wir L eine **Sprache** (über Σ).
Das neutrale Element von $(\Sigma^*, \cdot)$ nennen wir ϵ.

Satz:

Es seien Σ und Γ Alphabete. Jede Abbildung $\Sigma \rightarrow \Gamma^*$ lässt sich eindeutig auf einen Homomorphismus $\Sigma^* \rightarrow \Gamma^*$ erweitern.

Beweis:

Es sei h: $\Sigma^* \rightarrow \Gamma^*$ ein Homomorphismus. Dann ist $h(w) = h(w_1 \ldots w_n)$ mit $w_i \in \Sigma$
= $h(w_1) \ldots h(w_n)$, weil h homomorph.

□

Beispiel:

$$\Sigma = \{a, b\}$$

$$h(a) = b, \; h(b) = ab$$

$$h(abba) = bababb$$

Definitionen:

1. Die Länge |w| eine Wortes $w \in \Sigma^*$ ist n, falls $w = w_1 \ldots w_n$ mit $w_i \in \Sigma$.
2. Es seien $L_1, L_2 \subseteq \Sigma^*$ Sprachen über Σ. Dann ist $L_1 L_2 := \{uv \mid u \in L_1, v \in L_2\}$
3. Falls $L \subseteq \Sigma^*$, dann ist: $L^0 := \{\epsilon\}$, $L^1 := L$, $L^n := L L^{n-1}$ für $n \geq 1$
$$L^* := \bigcup_{n \geq 0} L^n.$$
   L^* wird die **Kleene´sche** Hülle genannt.

Beispiel:

1.) $\emptyset^* = \{\epsilon\}$
2.) $\emptyset \{ab, \epsilon\} = \emptyset$

1.2 Reguläre Ausdrücke

15.04.2005

Ein regulärer Ausdruck ist eine kurze, prägnante Schreibweise für „einfache“ Sprachen. Reguläre Ausdrücke eignen sich besonders für Ein-/Ausgabe.

Definition:

Wir definieren reguläre Ausdrücke über Σ induktiv:
Es sei Σ ein Alphabet.

1. $\emptyset$ ist ein regulärer Ausdruck.
2. ϵ ist ein regulärer Ausdruck.
3. a ist ein regulärer Ausdruck für $a \in \Sigma$.
4. rs ist ein regulärer Ausdruck, falls r, s reguläre Ausdrücke sind.
5. $r+s$ ist ein regulärer Ausdruck, falls r, s reguläre Ausdrücke sind.
6. r^* ist ein regulärer Ausdruck, falls r ein regulärer Ausdruck ist.

Wir ordnen jedem regulären Ausdruck r eine **Sprache** L(r) zu:

1. $L(\emptyset)=\emptyset$
2. $L(\epsilon)=\{\epsilon\}$
3. $L(a)=\{a\}$
4. $L(rs)=L(r)L(s)$
5. $L(r+s)=L(r)\cup L(s)$
6. $L(r^*)=L(r)^*$

Beispiele:

- $(A+...+Z+a+...+z)^*\ (O+o)\ t\ t\ o(A+...Z+a+...+z)^*$
- $L(0^*(10^*1)^*0^*)=0^*(10^*1)^*0^*$ sind alle Wörter über {0,1} mit gerade Anzahl von 1en.
- Wörter, die nicht 01 als Unterwort enthalten: 1^*0^*
- $(0+11+(10(1+00)^*01))^*$: Binärzahlen, die durch 3 teilbar sind.

19.04.2005

Definition:

$r^+ := rr^*$

$L^+ := \bigcup_{n>0} L^n = L L^*$

Bemerkung: $\emptyset^* = \{\epsilon\}$

Beispiel:

- $\epsilon \emptyset = \emptyset$
- $(a+b)^* = (a^* b^*)^*$

Satz:

Es seien $A, B, C \subseteq \Sigma^*$.

1.) $A(BC)=(AB)C$
2.) $\{\epsilon\} A = A\{\epsilon\} = A$
3.) $(A^*)^* = A^*$
4.) $A(B \cup C) = AB \cup AC$
5.) $(A \cup B)C = AC \cup BC$

Definition:

Eine Sprache $L \subseteq \Sigma^*$ ist **regulär**, falls es einen regulären Ausdruck r über Σ gibt mit L(r) = L.
Die Klasse der regulären Sprachen besteht aus allen regulären Sprachen über alle Alphabete.

Satz:

Reguläre Sprachen sind unter Vereinigung, Konkatenation, Kleene´sche Hülle und Homomorphismus abgeschlossen.
Das heißt:
Falls A,B regulär sind, dann auch

- $A \cup B$
- A^*
- AB
- $h(A)$ für Homomorphismus $h: \Sigma^* \to \Gamma^*$

Beweis:

A,B regulär => $A = L(r_A),\ B = L(r_B)$
$A \cup B = L(r_A + r_B)$
$AB = L(r_A r_B)$
$A^* = L(r_A)^*$

Nun zu h:

Fall 1: $r_A = \emptyset$ => $A = \emptyset$ => $h(A) = \emptyset$
Fall 2: $r_A = \epsilon$ => $A = \{\epsilon\}$ => $h(A) = \{\epsilon\}$
Fall 3: $r_A = a \in \Sigma$ => $A = \{a\}$ => $h(A) = \{b_1 \ldots b_n\}$
Fall 4: $r_A = r_B + r_C$ => $h(A) = h(B) \cup h(C)$. IV => $h(B), h(C)$ regulär.
Fall 5: ...
Fall 6: ...

□

Bemerkung: Abschluss unter Schnitt lässt sich so nicht beweisen.

Fragestellungen:

1.Wortproblem:

Eingabe: reguläre Sprache L, Wort w
Frage: Ist $w \in L$?

2.Sprachäquivalenz:

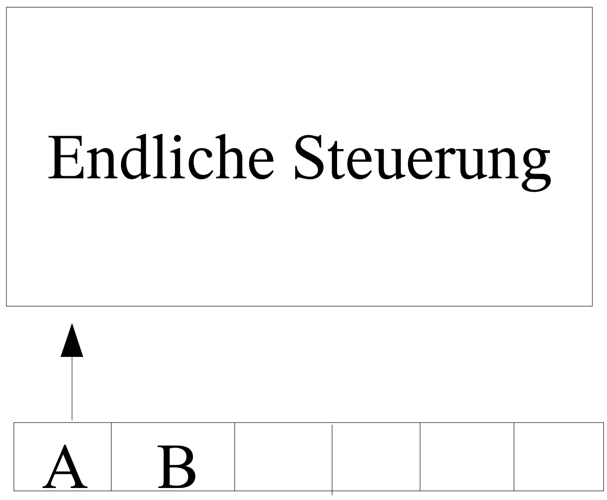
Eingabe: reguläre Sprachen L_1, L_2
Frage: $L_1 = L_2$?

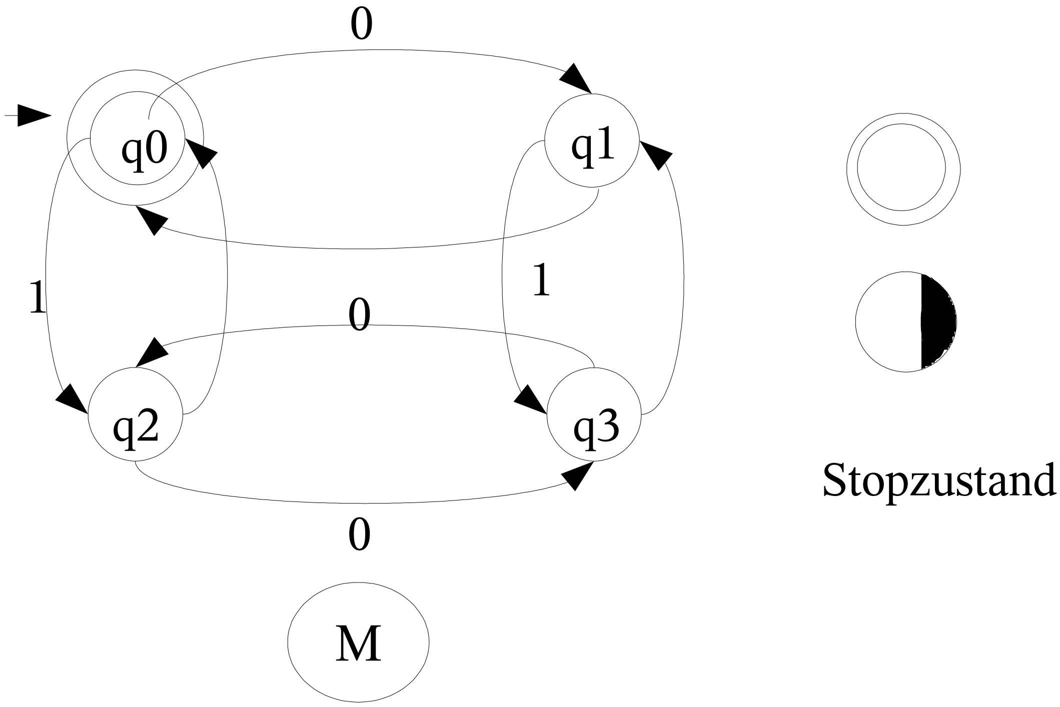
3.Schnittleerheitsproblem:

Eingabe: reguläre Sprachen L_1, L_2
Frage: $L_1 \cap L_2 = \emptyset$?

2 Automatentheorie

2.1 Endliche Automaten





Definition:

Eine (deterministischer) endlicher Automat (DFA) ist ein 5 – Tupel
$M=(Q, \Sigma, \delta, q_0, F)$,wobei

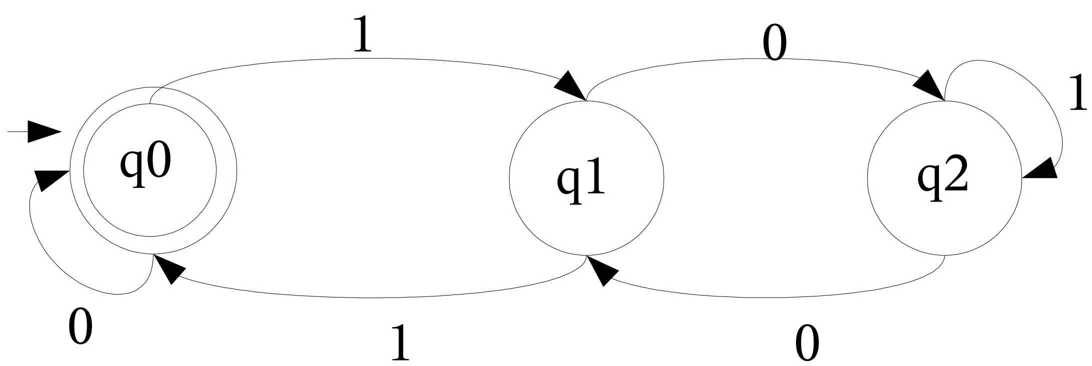
- Q endliche Menge von **Zuständen**
- Σ Eingabealphabet
- $\delta : Q\, x\, \Sigma \rightarrow Q$ Übergangsfunktion
- $q_0 \in Q$ **Startzustand**
- $F \subseteq Q$ Endzustände

Erweitere δ zu $\hat{\delta} : Q\, x\, \Sigma \rightarrow Q$ mittels

- $\hat{\delta}(q, \epsilon) := q$
- $\hat{\delta}(q, a\omega) := \hat{\delta}(\delta(q, a), \omega)$

Die Sprache, die M akzeptiert, ist
$L(M) := \{\omega \in \Sigma^* | \hat{\delta}(q_{0,} \omega) \in F\}$

Beispiel:



$M=(Q, \Sigma, \delta, q_{0,} F)$
$Q=\{q_0, q_1, q_2\}$
$\Sigma=\{0,1\}$
$\delta : Q\, x\, \Sigma \rightarrow Q$
$(q_0, 0) \rightarrow q_0$, $(q_0, 1) \rightarrow q_1$
$(q_1, 0) \rightarrow q_2$, $(q_1, 1) \rightarrow q_0$
$(q_2, 0) \rightarrow q_1$, $(q_2, 1) \rightarrow q_2$
$F=\{q_0\}$