# Automaten und Formale Sprachen

## SoSe 2013 in Trier

Henning Fernau
Universität Trier
fernau@uni-trier.de

19. April 2013

# Organisatorisches

**Vorlesung** FR 10.15-11.45 im F 59
Verschiebungsangebot: FR 10.10-11.40

**Übungsbetrieb** in Form von einer Übungsgruppe
BEGINN: in der zweiten Semesterwoche MI 09.10-09.55 im HS 13
Ab dann finden sie (im Wesentlichen) **vierzehntägig** statt.
Also ist die zweite Übung am 08.05., 08.25-9.55 im HS 13.
Bitte beachten Sie Stud.IP bzw. und unsere Institutsseite.

**Dozentensprechstunde** DO 14-15 in meinem Büro H 410 (4. Stock)

**Mitarbeitersprechstunde** (Markus Schmid) jederzeit

**Tutorensprechstunde** TBA H 412

**Zulassungskriterien** AFS gehört zusammen mit der “Schwesterveranstaltung” Berechenbarkeit und Komplexität im Wintersemester zu einem Modul.

Um zur Modulprüfung zugelassen zu werden, muss man für den AFS-Teil folgende Leistungen erbringen:

- 40% der Übungsaufgabenpunkte
- Bearbeitung jeder Aufgabe
- Wird eine Aufgabe eines Übungsblattes nicht bearbeitet, so wird das gesamte Blatt mit 0 Punkten bewertet.

## Bearbeitung einer Aufgabe

Eine Aufgabe kann auf zweierlei Art bearbeitet werden:

1. Sie präsentieren eine Lösung der Aufgabe (wie üblich).
2. Sie formulieren wenigstens eine Frage oder Aussage, die erklärt, weshalb Sie keine Lösung präsentiert haben oder präsentieren konnten.
   Dieser Kommentar muss sich auf den Inhalt der Vorlesung oder Aufgabenstellung beziehen und sollte nicht bloß allgemeiner Natur sein.

## Abgabe von Übungsaufgaben

- Abgaben sollten einzeln erfolgen.
- Abgabe 1: Kommentare zu einzelnen Aufgaben bis DI 7:30 Uhr.
- Abgabe 2: Lösungspräsentationen zu den (übrigen) Aufgaben: bis MI 8 Uhr.
- Abgabeort: mit AFS beschrifteter Kasten im 4. Stock vor dem Sekretariat von Prof. Näher. Alternativ unmittelbar vor der anschließenden Übung.
- Es dürfen auch Kommentare zu Aufgaben abgegeben werden, zu denen auch noch eine Lösung präsentiert wird.
- Es dürfe darüber hinaus auch Kommentare zur Übungs- oder Vorlesungsteilen abgegeben werden, die nicht unmittelbar in den Übungen abgefragt werden.

## Weitere Regeln

- Verspätete Abgaben gelten als nicht abgegeben und werden dementsprechend mit 0 Punkten bewertet.

- Die Lösungen sind handschriftlich anzufertigen; weder Schreibmaschinen- noch Computerausdrucke werden akzeptiert, erst recht keine Kopien.

- In der nächsten Übung werden die korrigierten und "bepunkteten" Übungsaufgaben wieder zurückgegeben (in den Übungen).

- Lösungen sind immer ausführlich zu erläutern.

## Zum Übungsbetrieb

- Durch die abgegebenen Kommentare steuern Sie wesentliche Teile des Übungsverlaufs.
- Darin erklären Sie nämlich, welche Teile der Vorlesung näher erläutert werden müssen.
- Aufgabe der Übungen oder des Übungsleiters ist es nicht, die gestellten Aufgaben der Reihe nach vorzurechnen oder Musterlösungen herauszugeben.
- Musterlösungen helfen auch nur eingeschränkt für Prüfungsvorbereitungen.

## Wo finde ich was bei AFS?

Vorlesungen (Foliensätze) liegen zum Runterladen **auf unserer Institutsseite** bereit.

Übungen finden Sie unter Stud.IP bzw. unserer Institutsseite.

Bitte melden Sie sich **sowohl für die VL als auch für die Übungen** im LSF-System an.

## Probleme ? Fragen ?

Klären Sie bitte Schwierigkeiten mit Vorlesungen oder Übungen möglichst umgehend in den zur Verfügung gestellten Sprechzeiten.

In der Tutorensprechstunde steht Ihnen Xenia Klinge zu Rückfragen bereit.

Wir helfen Ihnen gerne!
. . . wir sind aber keine Hellseher, die Ihnen Ihre Schwierigkeiten an der Nasenspitze ansehen. . .

**Modulprüfungen**
werden bei mir

- in der ersten Runde als schriftliche Prüfungen, aber

- in der zweiten Runde (“Nachprüfungsrunde”) als mündliche Prüfungen abgelegt.

Erscheinen Ihnen die Modulprüfungen am Ersttermin als zu gedrängt, so können Sie auch die zweite Prüfungsrunde zur Erstprüfung nutzen!

## Automaten und Formale Sprachen

Gesamtübersicht

- Organisatorisches
- Einführung
- Endliche Automaten und reguläre Sprachen
- Kontextfreie Grammatiken und kontextfreie Sprachen
- Chomsky-Hierarchie

## Einführung 1: Einordnung

- AFS und “Berechenbarkeit und Komplexität” (BK) (sowie teilweise Diskrete Strukturen und Logik) werden mancherorts zu einer oder zwei “4+2-Vorlesungen” “Einführung in die Theoretische Informatik” zusammengefasst.

- AFS liefert Grundlagen für bzw. ist verwandt mit:
  Compilerbau
  Textverarbeitungsalgorithmen
  computergestützte Linguistik / linguistische Datenverarbeitung
  Schaltkreisentwurf / Hardwarebeschreibung
  allgemein: formale Methoden in Spezifikation
  formale Beschreibung von Algorithmen, z.B. VL “Lernalgorithmen”

## Einführung 2: Motivation

AFS und BK untersucht die Frage:

Was ist die Natur einer Berechnung / eines Algorithmus ?

oder

Was ist ein “Rechner” (Computer) als mathematisches Objekt ?

Die Vorlesungen beschäftigen sich also mit den **absoluten Grundbegriffen der Informatik** als “Computer Science.”
Hoffnung: Es müsen Fragen der folgenden Art beantwortet werden:

- Was können Rechner prinzipiell (Berechenbarkeitsfragen) ?
- Was können Rechner “effizient” (d.h. schnell oder mit “sparsamem Speicher”) (AFS und Komplexitätstheorie) ?
- Wie können Rechner mathematisch modelliert werden ?
- Wie können Algorithmen in einer Weise notiert werden, dass man über sie mathematisch argumentieren kann ? (Mein Programm ist besser, weil . . . )

## Einführung 3: Literatur

E. Kinber / C. Smith: Theory of Computing. A Gentle Introduction. Prentice Hall.

U. Schöning: Theoretische Informatik kurz gefasst. BI / Spektrum.

Es gibt zahlreiche gute Skripten im Internet, z.B.: Skripten zur Vorlesung Informatik-B2 an der Universität Duisburg-Essen (Prof. Luther, Prof. Hertling)

Wie Sie auch bei anderen Veranstaltungen lernen werden, unterscheiden sich die Formalisierungen oft (leider) in manchen Einzelheiten von Buch zu Buch; am besten suchen Sie nach dem Buch, das Ihrem Geschmack am nächsten kommt. Ich werde mich vornehmlich am Kinber / Smith orientieren.

## Einführung 4: Was ist eine Sprache ?

. . . fast schon eine linguistische bis philosophische Frage . . .

Was würden Sie sagen?

## Einführung 4: Was ist eine Sprache ?

. . . fast schon eine linguistische bis philosophische Frage . . .

Eine Sprache ist eine Ansammlung von Elementen wie:

- Phonemen, Wortbestandteilen
- Wörtern, Satzbestandteilen
- Sätzen
- Texten
- Wichtig zum Lernen: Aussprache, Wörterbücher, Grammatik

## . . . und genauer?

Eine Menge $\Sigma$ heißt *Alphabet*, falls $\Sigma$ eine endliche, nicht-leere Menge ist, i.Z.: $|\Sigma| < \infty$ und $\Sigma \neq \emptyset$.
Die Elemente eines Alphabetes heißen *Buchstaben* oder *Zeichen*.

Man findet häufig folgende kurze Definition in der Literatur:
Eine *Sprache* $L$ (über $\Sigma$) ist eine Teilmenge des von der Menge $\Sigma$ frei erzeugten Monoids, i.Z.: $L \subseteq \Sigma^*$. Die Elemente einer Sprache heißen auch *Wörter*.

Alles klar ?!

Hoffentlich ist klar, was ein Alphabet sein soll.
Bsp.: $\Sigma = \{a, b, c, d, e, f, g, \ldots, y, z\}$

## Wörter und Sprachen

Wir haben verstanden, dass (abstrakt) eine Sprache eine Menge von Wörtern ist. (Linguisten nennen diese “Wörter” übrigens gerne “Sätze”.)

Wir müssen noch verstehen, was ein Wort sein soll.

Umgangssprachlich könnten wir sagen:
*Ein Wort ist eine Aneinanderreihung von Buchstaben.*
**Daraus folgt**: Reiht man zwei Wörter aneinander, so ergibt sich wieder ein Wort.

So wäre *ichliebeafs* ein Wort über dem Alphabet $\Sigma = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, e, \mathrm{f}, \mathrm{g}, \ldots, \mathrm{y}, z\}$.
(Vielleicht bräuchten wir noch Leerzeichen und Satzzeichen in unserem Alphabet.)

Aber was bedeutet “aneinanderreihen” genauer?
Wie können wir das mathematisch modellieren?
Eine endliche **Folge** von Buchstaben. $\rightsquigarrow$ DS-Wiederholung

## Zur Modellierungsfrage

Was wären denn alles Wörter in unserer Formalisierung?
Entspricht das unserer Intuition bzw. unserem Modellierungsvorhaben?

- “Normale” deutsche Wörter sind Wörter über dem Alphabet

$$\Sigma_d = \{a, b, c, \ldots, z, ä, ö, ü, ß\} \cup \{A, B, C, \ldots, Z, Ä, O, Ü\}$$

- Deutsche Texte sind Wörter über dem Alphabet

$$\Sigma_{d-T} = \Sigma_d \cup \{,\} \cup \{.\} \cup \{;\} \cup \{!\} \cup \{?\} \cup \{\_\}$$

  Das letzte Symbol soll das Leerzeichen andeuten.
  Mögliche weitere Symbole: Gedankenstrich, Anführungszeichen, ...
  In gedruckten Texten finden wir auch noch z.B. Ligaturen.

- Programmtexte, ...

## **Kartesische Mengenprodukte** (Wiederholung DS)

Es seien $X, Y$ Mengen.
Das *kartesische Produkt* oder *Mengenprodukt* von $X$ und $Y$ beinhaltet diejenigen *geordneten Paare* $(x, y)$ mit $x \in X$ und $y \in Y$, i.Z.:

$$X \times Y = \{(x, y) \mid x \in X \wedge y \in Y\}.$$

Lemma: Gilt $|X|, |Y| < \infty$, so gilt $|X \times Y| = |X| \cdot |Y|$.

Spezialfall: $X^2 = X \times X$.

Beispiel: $X = \mathbb{R}$, d.h.: $\mathbb{R}^2$ beschreibt die kartesische *Ebene*.

## Mengenpotenzen allgemein

Ist $X$ eine Menge und $n \in \mathbb{N}$, $n > 1$, so definiere: $X^1 := X$ und $X^n = X \times X^{n-1}$.
Dies ist ein Beispiel für eine *induktive Definition*.

Def.: Ein Element aus $X^n$ heißt auch *Folge der Länge $n$ über $X$*.
**Ist $X$ ein Alphabet, so nennen wir eine Folge auch ein *Wort* (der Länge $n$).**

Lemma: Gilt $|X| < \infty$, so ist $|X^n| = |X|^n$ für beliebige $n \in \mathbb{N}, n \geq 1$.

## Verkettung: eine Verknüpfung auf $X^n$ ?

Beispiel: Nach der rekursiven Definition z.B. für $\{a, b, c\}^3$ gilt:

$$(a, (a, b)) \in \{a, b, c\}^3.$$

Die Folge $(a, (a, b))$ der Länge drei entsteht durch *Verkettung* (oder *Hintereinanderschreiben*, *(Kon-)Katenation*) des Wortes $a$ der Länge eins mit dem Wort $(a, b)$ der Länge zwei, und letzteres wieder durch Verkettung der Wörter $a$ und $b$ der Länge eins.

Problem dieser Modellierung des Begriffs "Wort": Reiht man zwei Wörter (der Länge $n$) aneinander, so ergibt sich kein Wort der Länge $n$.
Mathematischer Grund: $X^n$ ist bezüglich der als $\cdot$ geschriebenen Operation "Verkettung" nicht abgeschlossen.

## Verkettung: eine Verknüpfung auf $X^+$ !

Lösung: Betrachte $X^+ := \bigcup_{n \geq 1} X^n$.

Lemma: Für alle $n \geq 2$ und alle $1 \leq \ell < n$ gibt es eine natürliche Bijektion zwischen $X^n$ und $X^{n-\ell} \times X^\ell$.
Das bedeutet: $X^n$ und $X^{n-\ell} \times X^\ell$ bezeichnen für jedes $1 \leq \ell < n$ "dasselbe."
Wir können also für alle $x_1, \ldots, x_n \in X$
$(x_1, \ldots, x_n) \in X^n$ mit $((x_1, \ldots, x_{n-\ell}), (x_{n-\ell+1}, \ldots, x_n))$ identifizieren.
Das beschreibt auch die "natürliche Bijektion" für das Lemma.

Mit dieser Identifikation gilt:
Satz: $(X^+, \cdot)$ ist eine Halbgruppe.
Deshalb (Assoziativität) kann man auch die vielen Klammern bei der Notation von Wörtern fortlassen: Wir schreiben also $aab$ statt $(a, (a, b))$.
Beispiel: $LASS \in \{A, D, S, L\}^4$ und $DAS \in \{A, D, S, L\}^3$, also gilt für die Konkatenation $LASSDAS \in \{A, D, S, L\}^7$.

**Halbgruppen** (Wiederholung DS)

Eine *Halbgruppe* ist eine Struktur $(H, \circ)$, wobei $\circ$ eine assoziative Verknüpfung auf $H$ ist.

Ein Element $e \in H$ heißt *neutrales Element* gdw. $x \circ e = e \circ x = x$ für alle $x \in H$.

Lemma: In einer Halbgruppe gibt es höchstens ein neutrales Element.

Beweis: Wenn die Aussage falsch wäre, so gäbe es zwei Elemente $e_1$ und $e_2$, die die Eigenschaft eines neutralen Elements erfüllen. Daher gilt:

$$e_1 = e_1 \circ e_2 = e_2 \circ e_1 = e_2.$$

Dies widerspricht der Annahme.

Satz: $(X^+, \cdot)$ ist eine Halbgruppe.
Was müssen wir also zeigen?

- Wie ist die Operation $\cdot : X^+ \times X^+ \to X^+$ genau definiert?
  Für $x, y \in X^+$ setze: $x \cdot y := (x, y)$. (Das ist nur sinnvoll mit dem Lemma, s.u.)

- Sind $x, y \in X^+$, so auch $x \cdot y \in X^+$ (Abgeschlossenheit).
  Da $x \in X^+$, gilt $x \in X^n$ für eine Zahl $n$; analog: $y \in X^m$ für ein $m$.
  Damit ist $(x, y) \in X^n \times X^m$, was wir mit $X^{n+m} \subseteq X^+$ identifizieren.

- Sind $x, y, z \in X^+$, so gilt: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
  Sei $x \in X^n$, $y \in X^m$ und $z \in X^r$. Also gibt es Buchstaben $a_1, \ldots, a_{n+m+r} \in X$ mit:
  $x = a_1 \cdots a_n$, $y = a_{n+1} \cdots a_{n+m}$, $z = a_{n+m+1} \cdots a_{n+m+r}$.
  Nach dem Lemma ist $x \cdot y = a_1 \cdots a_n \cdot a_{n+1} \cdots a_{n+m}$ "=" $a_1 \cdots a_{n+m} \in X^{n+m}$ und damit
  $(x \cdot y) \cdot z = a_1 \cdots a_{n+m} \cdot a_{n+m+1} \cdots a_{n+m+r}$ "=" $a_1 \cdots a_{n+m+r} \in X^{n+m+r}$.
  Analog sieht man ein: $x \cdot (y \cdot z)$ "=" $a_1 \cdots a_{n+m+r} \in X^{n+m+r}$. □

**Eigenschaften zweistelliger Operationen** hatten wir in DS studiert.

Wie sieht dies mit der Konkatenation aus?

Lemma: Es sei $X$ ein Alphabet mit $|X| > 1$.

Die Konkatenation auf $X^+$ ist **nicht** kommutativ.

Beweis: Betrachte zwei verschiedene Elemente $a, b$ aus $X$.

Dann gilt: $a \cdot b = (a, b) \neq (b, a) = b \cdot a$. □

Lemma: Die Konkatenation ist **nicht** idempotent.

Beweis: Es sei $X$ ein Alphabet. Betrachte einen bel. Buchstaben $a \in X$.

Da $X = X^1$, ist $a$ ein Wort der Länge 1. $a \cdot a \in X^2$ ist aber ein Wort der Länge 2. □

Ebenso sieht man, dass die Konkatenation (auf $X^+$) keine neutralen oder absorbierenden Elemente besitzt.

**Erzeugendensysteme** war ein weiterer Begriff aus DS.

Intuition: Eine Menge $X$ von Elementen der Grundmenge $H$ heißt Erzeugendensystem, wenn sich jedes Element aus $H$ durch Operationen der Algebra durch Elemente aus $X$ beschreiben lässt.

Offenbar lässt sich jedes Wort aus $X^+$ durch seine Buchstaben beschreiben (mit der Konkatenation als einziger Operation).

In diesem Sinne erzeugt $X$ die Halbgruppe $X^+$.

## Verkettung: eine Verknüpfung auf $X^*$ !

Problem: $X^+$ ist kein Monoid ?!

Lösung: Betrachte $X^* := \bigcup_{n \geq 0} X^n = X^+ \cup \{\lambda\}$. $\lambda$ (andere Notationen: $\epsilon$ oder $e$) ist *das leere Wort*, formal ein künstlich hinzugefügtes neutrales Element.

Satz: $(X^*, \cdot, \lambda)$ ist ein Monoid, das so genannte *frei erzeugte Monoid (über $X$)*.

Hinweis: Während "erzeugt" sich auf die Darstellbarkeit jedes Elements bezieht, bedeutet "Freiheit" die Eindeutigkeit der Darstellung.

**Was ist ein Monoid ?** (Wiederholung DS)

Eine Struktur $(M, \circ, e)$ heißt *Monoid* gdw.:

- $M$ ist eine Menge und $\circ$ ist eine zweistellige Operation (*Verknüpfung*) auf $M$, d.h., $\circ$ kann aufgefasst werden als Abbildung $\circ : M \times M \to M$ (*Abgeschlossenheit* von $M$ unter $\circ$).

- $\circ$ ist *assoziativ*, d.h.: für alle $x, y, z$ aus $M$ gilt: $(x \circ y) \circ z = x \circ (y \circ z)$, i.Z.:

$$\forall x, y, z \in M : (x \circ y) \circ z = x \circ (y \circ z).$$

- $e \in M$ ist *neutrales Element* von $\circ$, d.h.:

$$\forall x \in M : e \circ x = x \circ e = x.$$

## Beispiele

Es sei $\mathbb{N}$ die Menge der natürlichen Zahlen, d.h.:

$$\mathbb{N} = \{0, 1, 2, 3, 4, 5, \ldots\}.$$

Hinweis: $\mathbb{N}$ umfasst "im Wesentlichen" alle auf Rechnern tatsächlich behandelbaren Objekte; insbesondere "reelle Zahlen" sind Fiktion.

Beispiel: $(\mathbb{N}, \max, 0)$ ist ein Monoid.

Beispiel: $(\mathbb{N}, +, 0)$ ist ein Monoid.

Beispiel: $(\mathbb{N}, \min)$ ist eine Halbgruppe ohne neutrales Element.
Man könnte allerdings ein neutrales Element zu $\mathbb{N}$ künstlich hinzufügen.
Nennen wir es $\infty$, so bildet $(\mathbb{N} \cup \{\infty\}, \min, \infty)$ ein Monoid.

## Endliche Monoide

Ein Monoid $(M, \circ, e)$ heißt endlich gdw. $|M| < \infty$.

Endliche Monoide kann man gut mit einer *Verknüpfungstafel* angeben.

Beispiel: Betrachte $M = \mathbb{Z}_m = \{0, 1 \ldots, m-1\}$ mit der Addition $+$ modulo $m$ als Verknüpfung. Für $m = 3$ ergibt sich folgende Tafel:

| $+$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0 | 1 | 2 |
| 1 | 1 | 2 | 0 |
| 2 | 2 | 0 | 1 |

Konvention: “Zeile mal Spalte” (egal bei Kommutativität)

## Ablesen von Eigenschaften anhand einer Verknüpfungstafel

Da die mit der 0 indizierte Zeile bzw. Spalte die Grundelemente in der “richtigen Reihenfolge” aufzählt, ist $0$ das neutrale Element von $+$.

Die Abgeschlossenheit ist offenbar (warum?),

und die Assoziativität kann man durch erschöpfende Analyse aller Fälle nachrechnen, z.B.:

$$(1+2)+2 = 0+2 = 2 = 1+1 = 1+(2+2).$$

## **Mengenoperationen als Halbgruppenoperationen** (Wiederholung DS)

Es sei $X$ eine Menge. Dann bezeichnet $2^X$ *Potenzmenge* von $X$, d.i. die Menge der Teilmengen von $X$.

Beispiel: Ist $X = \{0, 1, 2\}$, so ist $2^X = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{0, 2\}, \{1, 2\}, X\}$.

Lemma: Gilt $|X| < \infty$, so ist $|2^X| = 2^{|X|}$.

Satz: Für jede Menge $X$ sind

$$(2^X, \cup, \emptyset)$$

und

$$(2^X, \cap, X)$$

Monoide.

## **Abbildungen** (Wiederholung DS)

Es seien $X, Y$ nicht-leere Mengen.
Eine *Abbildung* (oder *Funktion*) $f : X \to Y$ ist eine Vorschrift, die jedem Element aus $X$ höchstens ein Element aus $Y$ zuordnet.
In dieser Vorlesung werden Abbildungen stets *total* sein, d.h., jedem Element aus $X$ wird genau ein Element aus $Y$ zugeordnet.
In diesem Sinne ist dann $Y^X$ die Menge aller Abbildungen von $X$ nach $Y$.

Lemma: Gilt $|X|, |Y| < \infty$, so ist $|Y^X| = |Y|^{|X|}$.

Hinweis: Es macht oft Sinn, $n$ mit $\mathbb{Z}_n$ zu identifizieren...

**Hintereinanderausführung von Abbildungen** (Wiederholung DS)

Sind $f : X \to Y$ und $g : Y \to Z$ Abbildungen, so ordnet
ihre *Hintereinanderausführung* (oder *Komposition*) $f \circ g : X \to Z$
(beachte das "Vertauschen" von $g$ und $f$ in der Schreibweise)
einem $x \in X$ dasjenige Element $z \in Z$ zu, das sich durch $z = g(f(x))$ ergibt, i.Z.:
$f \circ g : X \to Z, x \mapsto g(f(x))$.

Satz: $(X^X, \circ, id_X)$ ist ein Monoid, wobei $id_X(x) = x$ für alle $x \in X$ gilt (*Identität*).

## **Morphismen I** (Wiederholung DS)

Allgemein bezeichnet ein *(Homo-)Morphismus* eine *strukturerhaltende Abbildung*.

Für Halbgruppen gilt daher: Sind $(H, \circ)$ und $(G, \square)$ Halbgruppen, so ist eine Abbildung $h : H \to G$ ein *(Halbgruppen-)Morphismus* gdw.

$$\forall x, y \in H : h(x \circ y) = h(x) \square h(y).$$

Satz: Sind $(H, \circ)$ und $(G, \square)$ Halbgruppen und $h : H \to G$ ein Morphismus, so ist $(\{h(x) \mid x \in H\}, \square)$ eine Halbgruppe. Besitzt $(H, \circ)$ darüber hinaus ein neutrales Element $e \in H$, so ist $h(e)$ neutrales Element von $(\{h(x) \mid x \in H\}, \square)$.

Stärkerer Begriff der Strukturerhaltung: **Isomorphismus**

## Morphismen II

Für Monoide gilt: Sind $(H, \circ, e)$ und $(G, \Box, 1)$ Monoide, so ist eine Abbildung $h : H \to G$ ein *(Monoid-)Morphismus* gdw.

$$\forall x, y \in H : h(x \circ y) = h(x) \Box h(y)$$

sowie $h(e) = 1$.

Es ist notwendig, die zweite Bedingung auch zu prüfen, wie folgendes Beispiel lehrt: Beispiel: Betrachte die Monoide $(\mathbb{Z}_k, \min, k-1)$ für $k > 0$; die identischen Abbildungen $h_k : x \mapsto x$ können als Halbgruppenmorphismen $(\mathbb{Z}_k, \min) \to (\mathbb{Z}_{k+1}, \min)$ aufgefasst werden. Das neutrale Element $k-1$ von $(\mathbb{Z}_k, \min)$ wird aber nicht auf das neutrale Element $k$ von $(\mathbb{Z}_{k+1}, \min)$ abgebildet.

Satz: Sind $(H, \circ, e)$ und $(G, \Box, 1)$ Monoide und $h : H \to G$ ein Morphismus, so ist $(\{h(x) \mid x \in H\}, \Box, 1)$ ein Monoid.

## Morphismen III

Beispiel: $\ell$ ordne einem Wort $w \in \Sigma^*$ dasjenige $n$ zu, für welches $w \in \Sigma^n$ gilt. Die so definierte Abbildung $\ell : \Sigma^* \to \mathbb{N}$ ist ein Monoidmorphismus von $(\Sigma^*, \cdot, \lambda)$ nach $(\mathbb{N}, +, 0)$.

Beispiel: $(\mathrm{mod}\ m)$ ordnet einer natürlichen Zahl $n$ diejenige Zahl aus $\mathbb{Z}_m$ zu, die sich als Rest von $n$ beim Teilen durch $m$ ergibt. Für jedes $m \in \mathbb{N}, m > 0$, ist $(\mathrm{mod}\ m) : \mathbb{N} \to \mathbb{Z}_m$ ein Monoidmorphismus von $(\mathbb{N}, +, 0)$ nach $(\mathbb{Z}_m, +, 0)$.

Satz: Die Komposition von zwei Morphismen liefert einen Morphismus.

Beispiel: Für jedes $m \in \mathbb{N}, m > 0$, und jedes Alphabet $\Sigma$ ist $\ell_m := (\ell \circ (\mathrm{mod}\ m))$ mit $\ell_m(w) = \ell(w) \ (\mathrm{mod}\ m)$ ein Monoidmorphismus.

## Morphismen IV

Lemma: Sind $(M, \circ, e)$ und $(N, \square, 1)$ Monoide und ist $h : M \to N$ ein Morphismus, so gilt für jede Zahl $n \geq 2$ und
für alle Elemente $x_1, \ldots, x_n \in M$: $h(x_1 \circ \cdots \circ x_n) = h(x_1) \square \cdots \square h(x_n)$.

Beweis: Leichte Induktion zur Übung. $\square$

Satz: Ist $X$ ein Erzeugendensystem für ein Monoid $(M, \circ, e)$, und ist $(N, \square, 1)$ ebenfalls ein Monoid, so ist jeder Morphismus $h : M \to N$ bereits durch Angabe der Bilder $h(x)$ für jedes $x \in X$ vollständig beschrieben.

Beweis: Da $X$ Erzeugendensystem, lässt sich jedes Element $y \in M$ schreiben als $y = x_1 \circ \cdots \circ x_n$ für geeignet gewählte Elemente $x_1, \ldots, x_n \in X$. Definiere nun: $g(y) := h(x_1) \square \cdots \square h(x_n)$.
Da $h$ Morphismus, gilt mit dem Lemma $h(y) = g(y)$, woraus die Behauptung folgt. $\square$

## Reguläre Sprachen

Def.: Eine Sprache $L \subseteq \Sigma^*$ heißt *regulär* gdw. es ein endliches Monoid $(M, \circ, e)$, einen Monoidmorphismus $h : (\Sigma^*, \cdot, \lambda) \to (M, \circ, e)$ sowie eine endliche Menge $F \subseteq M$ gibt mit

$$L = \{w \in \Sigma^* \mid h(w) \in F\}.$$

Beispiel: Die Sprache $L = \{w \in \{a, b\}^* \mid \ell(w)$ ist gerade $\}$ ist regulär.

Beweis: Betrachte den Morphismus $\ell_2$ und die endliche Menge $\{0\} \subset \mathbb{Z}_2$.

## Die Gegenstände der Vorlesung

- Alphabet $\Sigma$.
- Wort der Länge $n$ (über $\Sigma$): $w \in \Sigma^n$.
- Wort (bel. Länge): $w \in \Sigma^*$.
- Sprache (über $\Sigma$): $L \subseteq \Sigma^*$.
- Sprachfamilie: Menge von Sprachen; Bsp.: die regulären Sprachen.

## Hinweise

Es bestehen starke Verknüpfungen zum vorausgegangenen DSL-Modul.
Der dort präsentierte Stoff wird nicht (mehr) im Einzelnen hier wiederholt.

Ein weiteres Beispiel:

Def.: Es seien $x, y$ Wörter über $\Sigma$. $x$ heißt *Anfangswort* oder *Präfix* von $y$, i.Z. $x \sqsubseteq y$, gdw. es gibt ein Wort $w \in \Sigma^*$ mit $x \cdot w = y$.

Satz: $\sqsubseteq$ ist eine Halbordnung auf $\Sigma^*$.

Wir werden uns bemühen, diese Bezüge aufzuzeigen aber weitgehend nicht vorauszusetzen.
Ebensowenig wird die hier gegebene Definition regulärer Sprachen die grundlegende Definition der Vorlesung sein.
Wie wir sehen werden, könn(t)en wir hierzu auch deterministische endliche Automaten verwenden, die einige von Ihnen aus der Schule kennen werden.
Die aufgezeigte formale Beschreibung der Menge aller Wörter ist jedoch grundlegend für folgende Beweisführungen, die meist auf dem Prinzip der Induktion beruhen werden.