

## Automatentheorie und Formale Sprachen

für die Studiengänge

- Angewandte Informatik
- Informatik - Technische Systeme

### 01 Begrüßung & Grundlagen

Prof. Dr. David Sabel  
Sommersemester 2025

Stand der Folien: 20. Mai 2025

## Personen

Josef Peter Heger (2 Übungen), Lehrbeauftragter

Prof. Dr. David Sabel (Vorlesung und 3 Übungen)

- Professor für Theoretische Informatik, Forschungsgebiet Programmiersprachensemantik
- Sprechstunde: nach Vereinbarung (Email schreiben)
- Büro: UDE C 031 (Haus C, Nord, EG) oder online
- Kontakt: David.Sabel@hs-rm.de bzw. www.davidsabel.de

## Vorlesung und Übungen

### Vorlesung

- Termin: Mittwoch, 08:15 - 09:45 Raum B002 (UDE)

### Übungen

- A: Montag 10:00-11:30 Uhr, UDE-C405 (Sabel)
- C: Dienstag 08:15-09:45 Uhr, UDE-C037 (Sabel)
- D: Dienstag 10:00-11:30 Uhr, UDE-C405 (Heger)
- B: Dienstag 11:45-13:15 Uhr, UDE-C407 (Sabel)
- E: Mittwoch 14:15-15:45 Uhr, UDE-C405 (Heger)

Beginn: 22.04.2025 (nächste Woche)

Feiertag: Bei Feiertag eine andere Übung derselben Woche besuchen!  
(Betrifft schon nächste Woche die Montagsübung!)

## Übungen

- Wöchentliche Aufgaben: selbstständig bearbeiten und abgeben
- Besprechung in den Übungen nach der Abgabe
- Nächste Woche: Präsenzaufgaben in der Übung (Aufgabenblatt 0), Aufgabenblatt 1 bis Sonntag 27.04. bearbeiten und abgeben
- Abgabe via sabato

## Material

- Skript, Folien, Übungsaufgaben
- Eventuell weiteres Material, Links, etc.
- Verfügbar in ILIAS:
  - Einloggen über StudIP → ILIAS
- Aufgaben und Abgabe in Subato
  - Teilweise Unit-Tests für Aufgaben



## Modulprüfung

### Studienleistung:

- 1 Fristgerechte Anmeldung in compass.hs-rm.de **Frist: 14.04.2025 bis 05.05.2025**
- 2 Fristgerechte Abgabe der bearbeiteten Aufgaben (mindestens 9 von 12 Blättern) ernsthafter Versuch, die Aufgaben zu bearbeiten
- 3 Lerngruppen sind gewünscht, aber Abgabe individuell aufschreiben und machen
- 4 Mindestens einmal eine Aufgabe in der Übung vorrechnen
- 5 Note durch (bestes) **benotetes Vorrechnen**

### Prüfungsleistung:

- 1 Anmeldung für die Klausur in compass.hs-rm.de Frist: 23.06.2025. bis 07.07.2025
- 2 Klausur (90 Minuten) im Prüfungszeitraum 21.07.2025 bis 22.08.2025

## Literatur: Wesentliche Quellen

- Vorlesungsskript
- Uwe Schöning: Theoretische Informatik - kurz gefasst, 5. Auflage, Spektrum Akademischer Verlag, 2008 (ältere Auflagen sind auch in Ordnung)
  - Teile sind u.U. zu kurz gefasst*

## Literatur: Weitere Literatur

- Alexander Asteroth und Christel Baier: Theoretische Informatik, Pearson Studium 2002.
  - Gutes Buch, Aufbau in anderer Reihenfolge*
- John E. Hopcroft, Rajeev Motwani und Jeffrey D. Ullman: Introduction to Automata Theory, Languages, and Computation, 3. Auflage, 2006
  - Der Klassiker, umfangreich (Erstauflage 1979!)*
- Ingo Wegener: Theoretische Informatik - eine algorithmenorientierte Einführung, 3. Auflage, Teubner Verlag, 2005.
  - Gutes Buch, andere Reihenfolge, Algorithmen stehen im Vordergrund.*
- Michael Sipser: Introduction to the Theory of Computation, 3. Auflage, 2021
  - Umfangreich und sehr gut*
- ...

# INHALTSÜBERSICHT

## Inhalte und Motivation der Veranstaltung

Drei wesentliche Themen der **Theoretischen Informatik**:

- 1 Formale Sprachen und Automatentheorie  
Fragestellungen und Anwendungen: *Wie stellt man Entscheidungsprobleme formal dar? Wie modelliere ich ein Problem? Wie spezifiziere ich meine Eingabesyntax? Wie erkenne ich die Eingabe anhand der Syntax? Welches Tool ist das richtige? Geht das überhaupt?, ...*
- 2 Berechenbarkeitstheorie  
Fragestellungen und Anwendungen: *Welche Probleme kann man algorithmisch (mit dem Computer) überhaupt lösen? Was sind die Konsequenzen für die Praxis?*
- 3 Komplexitätstheorie  
Fragestellungen und Anwendungen: *Welche Probleme kann man in annehmbarer Zeit lösen? Wie gehe ich mit den schwer lösbaren Problemen in der Praxis um?*

Wir: Schwerpunkt auf 1, Einblick in 2 und 3.

## Inhalte: Formale Sprachen und Automatentheorie

- Chomsky-Grammatiken und Chomsky-Hierarchie
- Reguläre Sprachen: reguläre Grammatiken, endliche Automaten, reguläre Ausdrücke, Äquivalenz der Formalismen, Pumpinglemma, Minimalautomaten, Abschlusseigenschaften
- Kontextfreie Sprachen: kontextfreie Grammatiken, Normalformen, Pumpinglemma, CYK-Algorithmus, Kellerautomaten, Abschlusseigenschaften
- Kontextsensitive Sprachen und Typ 0-Sprachen: kontextsensitive Grammatiken, Turingmaschinen

## Inhalte: Berechenbarkeit und Komplexitätstheorie

### Berechenbarkeit

- Intuitive Berechenbarkeit, Churchsche These
- Turing-Berechenbarkeit
- Halteproblem, Unentscheidbarkeit, Reduktionen
- Rekursiv aufzählbar

### Komplexitätstheorie

- Zeitkomplexität, Klassen P und NP
- NP-Schwere, NP-Vollständigkeit
- polynomielle Reduktionen
- das SAT-Problem und der Satz von Cook
- weitere NP-vollständige Probleme

# GRUNDLAGEN: WÖRTER UND FORMALE SPRACHEN

## Wörter

### Alphabet

Ein **Alphabet**  $\Sigma$  ist eine endliche nicht-leere Menge von **Zeichen** (oder **Symbolen**).

Z.B.  $\Sigma = \{a, b, c, d, e\}$

### Wort

Ein **Wort**  $w$  über  $\Sigma$  ist eine endliche Folge von Zeichen aus  $\Sigma$ .

Beispiele:

- *bade* ist ein Wort über  $\{a, b, c, d, e\}$
- *baden* ist **kein** Wort über  $\{a, b, c, d, e\}$

## Weitere Notationen zu Wörtern

- Das **leere Wort** wird als  $\varepsilon$  notiert.
- Für  $w = a_1 \cdots a_n$  ist  $|w| = n$  die **Länge des Wortes**
- Für  $1 \leq i \leq |w|$  ist  $w[i]$  das Zeichen an  $i$ . Position in  $w$ .
- Für  $a \in \Sigma$  und  $w$  ein Wort über  $\Sigma$  sei  $\#_a(w) \in \mathbb{N}_0$  die **Anzahl an Vorkommen des Zeichens  $a$**  im Wort  $w$

Beispiele:

- Es gilt  $|\varepsilon| = 0$  und  $\#_a(\varepsilon) = 0$  für alle  $a \in \Sigma$ .
- Für  $\Sigma = \{a, b, c\}$  ist
  - $|abbccc| = 6$
  - $|aabbccc| = 8$
  - $\#_a(abbccc) = 1$
  - $\#_c(aabbccc) = 3$
- Für  $w = abbc d$  ist  $w[1] = a$ ,  $w[5] = c$  und  $w[7]$  undefiniert.

## Konkatenation und Kleene-Stern

### Konkatenation

Das Wort  $uv$  (alternativ  $u \circ v$ ) entsteht, indem Wort  $v$  hinten an Wort  $u$  angehängt wird.

$\Sigma^*$  bezeichnet die Menge aller Wörter über  $\Sigma$ :

### Definition von $\Sigma^i, \Sigma^*, \Sigma^+$

Sei  $\Sigma$  ein Alphabet, dann definieren wir:

$$\begin{aligned} \Sigma^0 &:= \{\varepsilon\} \\ \Sigma^i &:= \{aw \mid a \in \Sigma, w \in \Sigma^{i-1}\} \text{ für } i > 0 \\ \Sigma^* &:= \bigcup_{i \in \mathbb{N}_0} \Sigma^i \\ \Sigma^+ &:= \bigcup_{i \in \mathbb{N}} \Sigma^i \end{aligned}$$

Beachte:  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  und  $\mathbb{N} = \{1, 2, \dots\}$

## Beispiele

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

Dann ist

$$\Sigma^0 = \{\varepsilon\}$$

$$\Sigma^1 = \Sigma = \{a, b\}$$

$$\Sigma^2 = \{aa, ab, ba, bb\}$$

$$\Sigma^3 = \{xw \mid x \in \{a, b\}, w \in \Sigma^2\} = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$$

...

und

$$\Sigma^* = \{\varepsilon, a, b, aa, ab, ba, bb, aaa, aab, aba, abb, baa, bab, bba, bbb, aaaa, \dots\}$$

## Weitere Notationen und Begriffe

Sei  $w$  ein Wort über  $\Sigma$

- $w^m$  entsteht aus  $m$ -maligen Konkatenieren von  $w$ , d.h.

$$w^0 = \varepsilon \text{ und } w^m = ww^{m-1} \text{ für } m > 0$$

- $\bar{w}$  (manchmal auch  $w^R$ ) ist das rückwärts gelesene Wort  $w$ , d.h.

$$\bar{\varepsilon} = \varepsilon \text{ und für } w = a_1 \cdots a_n \text{ ist } \bar{w} = a_n a_{n-1} \cdots a_1$$

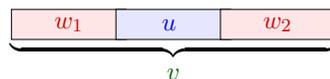
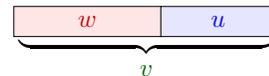
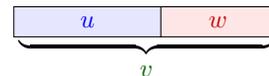
- $w$  ist ein Palindrom g.d.w.  $w = \bar{w}$

Beispiele für Palindrome: anna, reliefpfeiler, lagerregal, annasusanna

## Sprechweisen: Präfix, Suffix, Teilwort

Seien  $u, v$  Wörter über einem Alphabet  $\Sigma$ .

- $u$  ist ein **Präfix** von  $v$ , wenn es ein Wort  $w$  gibt mit  $uw = v$ .
- $u$  ist ein **Suffix** von  $v$ , wenn es ein Wort  $w$  gibt mit  $wu = v$ .
- $u$  ist ein **Teilwort** von  $v$ , wenn es Wörter  $w_1, w_2$  gibt mit  $w_1uw_2 = v$ .



Beispiel: Sei  $w = ababbaba$

- $aba$  ist ein Präfix, Suffix und Teilwort von  $w$
- $ababb$  ist ein Präfix (und Teilwort) von  $w$ , aber kein Suffix von  $w$
- $bab$  ist Teilwort von  $w$ , aber weder ein Präfix noch ein Suffix

## Formale Sprache

### Formale Sprache

Eine (formale) Sprache  $L$  über dem Alphabet  $\Sigma$  ist eine Teilmenge von  $\Sigma^*$  d.h.  $L \subseteq \Sigma^*$

Beachte: Wir verwenden  $L$  für „language“.

### Operationen auf formalen Sprachen

Seien  $L, L_1, L_2$  formale Sprachen über  $\Sigma$

- **Vereinigung:**  $L_1 \cup L_2 := \{w \mid w \in L_1 \text{ oder } w \in L_2\}$
- **Schnitt:**  $L_1 \cap L_2 := \{w \mid w \in L_1 \text{ und } w \in L_2\}$
- **Komplement zu  $L$ :**  $\bar{L} := \Sigma^* \setminus L$
- **Produkt:**  $L_1 L_2 = L_1 \circ L_2 = \{uv \mid u \in L_1 \text{ und } v \in L_2\}$

## Quiz 1

Sei  $\Sigma = \{a, b\}$  und  $L_1 = \{a^i \mid i \in \mathbb{N}_0\}$  und  
 $L_2 = \{b^i \mid i \in \mathbb{N}_0\}$ .

Welche Sprache wird durch  $L_1 \cap L_2$  beschrieben?

- a)  $\emptyset$
- b) Sprache aller Wörter aus  $a$ 's und  $b$ 's
- c)  $\{\varepsilon\}$
- d) Sprache aller Wörter mit gleich vielen  $a$ 's und  $b$ 's

arsnova.hs-rm.de  
6750 1376



## Quiz 2

Sei  $\Sigma = \{a, b\}$  und  $L_1 = \{a^i \mid i \in \mathbb{N}_0\}$  und  
 $L_2 = \{b^i \mid i \in \mathbb{N}_0\}$ .

Welche Sprache wird durch  $\overline{L_1}$  beschrieben?

- a)  $\emptyset$
- b) Sprache aller Wörter, die mindestens ein  $b$  enthalten.
- c) Sprache aller Wörter, die mindestens ein  $a$  und ein  $b$  enthalten.
- d) Sprache aller Wörter, die kein  $a$  enthalten.

arsnova.hs-rm.de  
6750 1376



## Weitere Beispiele

Sei  $\Sigma = \{a, b\}$  und  $L_1 = \{a^i \mid i \in \mathbb{N}_0\}$  und  $L_2 = \{b^i \mid i \in \mathbb{N}_0\}$ .

- $L_1 \cup L_2 =$  Sprache aller Wörter, die nur aus  $a$ 's oder nur aus  $b$ 's bestehen
- $L_1 L_2 = \{a^i b^j \mid i, j \in \mathbb{N}_0\}$
- $L_1 L_1 = L_1$

## Kleenescher Abschluss

Sei  $L$  eine Sprache. Dann ist:

$$L^0 := \{\varepsilon\}$$

$$L^i := L \circ L^{i-1} \text{ für } i > 0$$

$$L^* := \bigcup_{i \in \mathbb{N}_0} L^i$$

$$L^+ := \bigcup_{i \in \mathbb{N}} L^i$$

Die Sprache  $L^*$  nennt man auch den **Kleeneschen Abschluss von  $L$**  benannt nach Stephen Cole Kleene (1909-1994).

Beispiel:  $L = \{ab, ac\}$

- $L^0 = \{\varepsilon\}$
- $L^1 = L \circ L^0 = L = \{ab, ac\}$
- $L^2 = \{abab, abac, acab, acac\}$
- $L^3 = \{ababab, ababac, abacab, abacac, acabab, acabac, acacab, acacac\}$
- $L^* = \{\varepsilon\} \cup \{ax_1 ax_2 \cdots ax_i \mid i \in \mathbb{N}, x_j \in \{b, c\}, j = 1, \dots, i\}$ .

### Quiz 3

Welche Sprache beschreibt

$\{0\} \cup (\{1, \dots, 9\} \circ \{0, \dots, 9\}^*)$ ?

arsnova.hs-rm.de  
6750 1376



### Quiz 4

Sei  $\Sigma = \{a, d, e, h, k, m, n, s, u, t, z\}$   
und  $L = \{hund, katze, maus\}$ .

Welche der Antworten ist richtig?

- a)  $katzemaushund \in L^2$
- b)  $haus \in L^*$
- c)  $tatze \in \Sigma^*$
- d)  $hundhundhund \in L^4$

arsnova.hs-rm.de  
6750 1376

