

FSK & TIMI

Begrüßung, Organisatorisches, Inhaltsübersicht und Grundlagen

Prof. Dr. David Sabel
LFE Theoretische Informatik



Letzte Änderung der Folien: 26. April 2022

Personen

- **Dozent:**
Prof. Dr. David Sabel
Email: david.sabel@ifi.lmu.de
- **Wissenschaftliche Mitarbeiter:innen:**
Sarah Vaupel
Stephan Barth
- **Tutor:inn:en und Korrektor:inn:en:**
Charlotte Gerhaher
David Mosbach
Elisabeth Lempa
Elisabeth Schwertfeller
Lea Korn
Luca Maio
Lukas Bartl
Michael Fink Amores
Simon Rossmair
Thomas Grill

Zielgruppe der Veranstaltung (Hörerkreis)

Formale Sprachen und Komplexität [FSK]:

- Studierende der Informatik
- Studierende der Bioinformatik
- Studierende im Lehramt
- Studierende im Nebenfach Informatik

Theoretische Informatik für Medieninformatiker [TIMI]:

- Studierende der Medieninformatik

Struktur der Veranstaltung

	Montag	Dienstag	Mittwoch	Donnerstag	Freitag
8-10	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; padding: 5px; background-color: #fff9c4;">Vorlesung M 018</div> <div style="border: 1px solid black; padding: 5px; background-color: #c8e6c9;">Zentralübung (A240) (optional, ca 2 wöchig)</div> </div> <p style="font-style: italic; font-size: 1.2em;">verschiedene Übungsgruppen (siehe uni2work)</p>				
10-12					
12-14					
14-16					
16-18					
18-20					

- **Vorlesung:** FSK: 3V, TIMI 2V (integriert, Plan auf Webseite)
- **Digitale Alternative:** ScreenCasts aus dem SoSe 2021
- **Zentralübung:** Zusatzangebot, Fragestunde & Beispiele (Plan auf Webseite) **Raum A 240**
- **Übungen:** präsenz oder online; Besprechung der Hausaufgaben; FSK: 2Ü, TIMI: 1Ü

Webseiten

Webseiten zu den Veranstaltungen:

www.tcs.ifi.lmu.de/lehre/ss-2022/fsk und www.tcs.ifi.lmu.de/lehre/ss-2022/timi

Anmeldung im Uni2Work:

- uni2work.ifi.lmu.de/course/S22/IfI/FSK
- uni2work.ifi.lmu.de/course/S22/IfI/TIMI

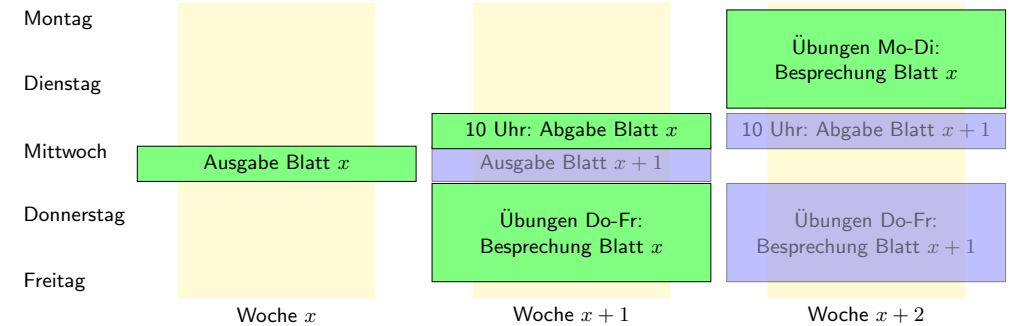
Anmeldung ist **notwendig** für:

- Zugriff auf Material, Abgabe & Korrektur der Hausaufgaben
- Anmeldung zu den Übungsterminen
- Anmeldung zur Prüfung (noch nicht online)

Zulip-Chat

- Server-Adresse: chat.ifi.lmu.de
 - Stream: **TCS-22S-FSK-TIMI**
- Fragen und Kommentare am besten dort stellen.

Hausaufgaben



- Übungen diese Woche ab Donnerstag: Kennenlernen+Besprechung Blatt 0 (ohne Abgabe)
- Ausgabe, Abgabe und Korrektur elektronisch über Uni2Work
- Prüfungsbonus für erfolgreiches Bearbeiten der Aufgaben

Korrektur und Bonuspunkte

- Ausgewählte Hausaufgaben werden bepunktet
- Für jede Lösung zu einer bepunkteten Aufgabe gibt es 0 oder 1 oder 2 Punkte

Bonusregelung (gilt für Prüfung und Nachholprüfung im SoSe 2022):

100% der erreichbaren Übungspunkte entsprechen 10% der Prüfungspunkte

$$\text{Prüfungsbonus} = \frac{\text{erreichte Übungspunkte}}{\text{maximale Übungspunkte}} \cdot 0,1 \cdot \text{maximale Prüfungspunkte}$$

wenn die Prüfung bestanden ist (Bonuspunkte **helfen nicht zum Bestehen**)

Die Prüfung ist auf jeden Fall bestanden, wenn 50% der Prüfungspunkte erreicht wurden.

Prüfungen

- **Plan (beantragt, noch nicht bestätigt):**
Erstklausur am 17.08.2022 und Nachklausur am 21.09.2022
- Anmeldung zur Prüfung wird noch freigeschaltet
- Bonuspunkte gelten für Prüfung und Nachprüfung
- Teilnahme an der Nachholprüfung auch ohne Teilnahme an der Prüfung möglich

Material

- Vorlesungsfolien
- Skript zur Vorlesung (wird nach und nach bereit gestellt):
Markierungen mit ★ für nicht-TIMI-relevante Teile
- ScreenCasts zur Vorlesung (aus SoSe 2021)
- Lehrbuch: Uwe Schöning, Theoretische Informatik – Kurz gefasst
- Hausaufgaben (Übungsblätter im Uni2Work)

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

Weitere Literatur:

- Alexander Asteroth und Christel Baier: Theoretische Informatik, Pearson Studium 2002.
Gutes Buch, Aufbau in anderer Reihenfolge, Zugriff über UB
- 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, Zugriff über UB

Inhaltsübersicht über die Veranstaltung

Drei große wesentliche Themen der **Theoretischen Informatik**:

- 1 Formale Sprachen und Automatentheorie
Wie stellt man Entscheidungsprobleme formal dar?
- 2 Berechenbarkeitstheorie
Welche Probleme kann man algorithmisch (bzw. mit dem Computer) überhaupt lösen?
- 3 Komplexitätstheorie
Welche Probleme kann man in annehmbarer Zeit lösen?

Inhalte: Formale Sprachen und Automatentheorie

- Chomsky-Grammatiken und Chomsky-Hierarchie
- Das Wortproblem und weitere Entscheidungsprobleme
- Reguläre Sprachen: reguläre Grammatiken, deterministische endliche Automaten, nichtdeterministische endliche Automaten, ϵ -Übergänge, reguläre Ausdrücke, Äquivalenz der Formalismen, Pumpinglemma, Satz von Myhill-Nerode, Minimalautomaten, Abschlusseigenschaften
- Kontextfreie Sprachen: kontextfreie Grammatiken, Chomsky-Normalform, Greibach-Normalform, Pumpinglemma, Ogden's Lemma, Cocke-Younger-Kasami-Algorithmus, Kellerautomaten, Abschlusseigenschaften
- Kontextsensitive Sprachen und Typ 0-Sprachen: kontextsensitive Grammatiken, Kuroda-Normalform, Turingmaschinen, Linear bounded automata (LBA), LBA-Probleme

TIMI: Zum Teil nur Auswahl der Inhalte / oberflächlichere Behandlung !

Inhalte: Berechenbarkeitstheorie

- Intuitive Berechenbarkeit, Churchsche These
- Turing-Berechenbarkeit, Varianten von Turingmaschinen (z.B. Mehrbandmaschinen)
- LOOP-, WHILE-, GOTO-Berechenbarkeit: LOOP-Programme, WHILE-Programme, GOTO-Programme, Äquivalenz zu Turingmaschinen
- Primitiv-rekursive Funktionen, Ackermannfunktion, μ -Rekursion
- Halteproblem, Unentscheidbarkeit
- Rekursiv aufzählbar
- Reduktionen
- Postsches Korrespondenzproblem

TIMI: Zum Teil nur Auswahl der Inhalte / oberflächlichere Behandlung !

Inhalte: Komplexitätstheorie

- Zeitkomplexität
- Klassen P und NP
- NP-Härte, NP-Vollständigkeit
- polynomielle Reduktionen
- das SAT-Problem
- Satz von Cook
- weitere NP-vollständige Probleme (z.B. 3-SAT, Clique, Vertex Cover, Subset Sum, Knapsack, Directed Hamiltonian Circuit, Hamiltonian Circuit, ...)

TIMI: Zum Teil nur Auswahl der Inhalte / oberflächlichere Behandlung !

Grundlagen: Worte und Formale Sprachen

Worte

Alphabet

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

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

Wort

Ein **Wort** w über Σ ist eine endliche Folge von Zeichen aus Σ .

Beispiele:

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

Weitere Notationen zu Worten

- Das **leere Wort** wird als ε 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 Σ sei $\#_a(w) \in \mathbb{N}$ 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 $w = abbcd$ 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.

Σ^* bezeichnet die Menge aller Wörter über Σ :

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

Sei Σ ein Alphabet, dann definieren wir:

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

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

Beispiele

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

Dann ist

$$\begin{aligned}\Sigma^0 &= \{\varepsilon\} \\ \Sigma^1 &= \Sigma = \{a, b\} \\ \Sigma^2 &= \{aa, ab, ba, bb\} \\ \Sigma^3 &= \{xw \mid x \in \{a, b\}, w \in \Sigma^2\} = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}\end{aligned}$$

...

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 Σ

- 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} 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

Formale Sprache

Formale Sprache

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

Beachte: Wir verwenden L für „language“.

Operationen auf formalen Sprachen

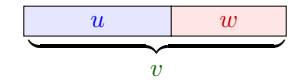
Seien L, L_1, L_2 formale Sprachen über Σ

- 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: $L_1 L_2 = L_1 \circ L_2 = \{uv \mid u \in L_1 \text{ und } v \in L_2\}$

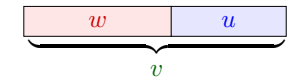
Sprechweisen: Präfix, Suffix, Teilwort

Seien u, v Wörter über einem Alphabet Σ .

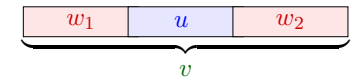
- u ist ein Präfix von v , wenn es ein Wort w gibt mit $uw = v$.



- 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 $w_1 u w_2 = v$.



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

Beispiele

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

- $L_1 \cup L_2 = ?$
- $L_1 \cap L_2 = ?$
- $\bar{L}_1 = ?$
- $L_1 L_2 = ?$
- $L_2 L_1 = ?$
- $L_1 L_1 = ?$

Für $L_1 = \{\spadesuit, \clubsuit, \diamondsuit, \heartsuit\}$ und $L_2 = \{7, 8, 9, 10, J, D, K, A\}$ stellt $L_1 L_2$ eine Repräsentation der Spielkarten eines Skatblatts dar.

Kleenescher Abschluss

Sei L eine Sprache. Dann ist:

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

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

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

$$L^+ := \bigcup_{i \in \mathbb{N}_{>0}} L^i$$

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

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

- $L^0 = \{\varepsilon\}$
- $L^1 = L \circ L^0 = L = \{ab, ac\}$
- $L^2 = \{abab, abac, acab, acac\}$
- $L^3 = \{ababab, ababac, abacab, abacac, acabab, acabac, acacab, acacac\}$
- $L^* = \{\varepsilon\} \cup \{ax_1ax_2 \cdots ax_i \mid i \in \mathbb{N}_{>0}, x_j \in \{b, c\}, j = 1, \dots, i\}$.

Weitere Beispiele

$$((\{\varepsilon, 1\} \circ \{0, \dots, 9\}) \cup (\{2\} \circ \{0, 1, 2, 3\})) \circ \{:\} \circ \{0, 1, 2, 3, 4, 5\} \circ \{0, \dots, 9\}$$

Beschriebene Sprache = ? **Sprache aller gültigen Uhrzeiten**

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

Beschriebene Sprache = ? **Sprache aller natürlichen Zahlen**