FSK & TIMI
Begrüßung, Organisatorisches, Inhaltsübersicht und Grundlagen

Prof. Dr. Jasmin Blanchette

Lehr- und Forschungseinheit für
Theoretische Informatik

Stand: 18. April 2023
Folien ursprünglich von PD Dr. David Sabel

Personen

- **Dozent:**
  Prof. Dr. Jasmin Blanchette
  Email: jasmin.blanchette@ifi.lmu.de
- **Wissenschaftlicher Mitarbeiter:**
  Jannis Limperg
- **Tutor*innen und Korrektor*innen:**
  Rafeek Dalate
  Fabian Grubmüller
  Manuel Hülskamp
  Vivian Kafadar
  Luca Maio
  Markus Sölderer
  Zhiyang Song

Zielgruppe der Veranstaltung (Hörerkreis)

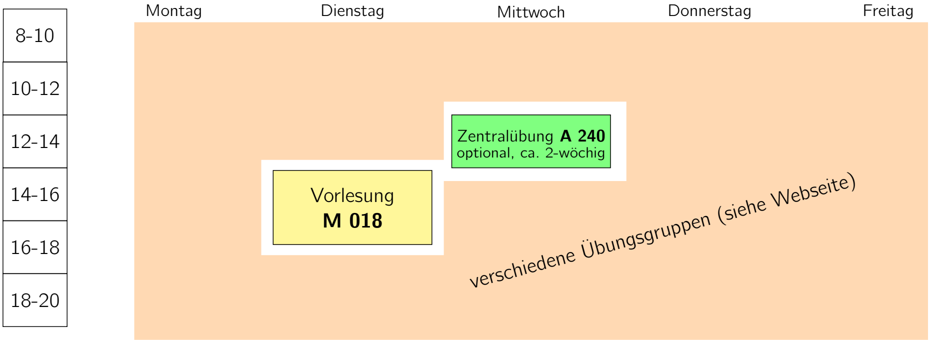
Formale Sprachen und Komplexität [FSK]:

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

Theoretische Informatik für Medieninformatiker [TIMI]:

- Studierende der Medieninformatik

Struktur der Veranstaltung



- **Vorlesung:** FSK: 3V, TIMI 2V (integriert, Plan auf Webseite)
- **Zentralübung:** Zusatzangebot, Fragestunde & Beispiele (Plan auf Webseite)
- **Übungen:** in Präsenz; Besprechung der Hausaufgaben; FSK: 2Ü, TIMI: 1Ü

Webseiten

Webseiten zu den Veranstaltungen:

www.tcs.ifi.lmu.de/lehre/ss-2023/fsk_de.html
www.tcs.ifi.lmu.de/lehre/ss-2023/timi_de.html

Anmeldung im Moodle:

- **moodle.lmu.de/course/view.php?id=27180**
- **moodle.lmu.de/course/view.php?id=27182**

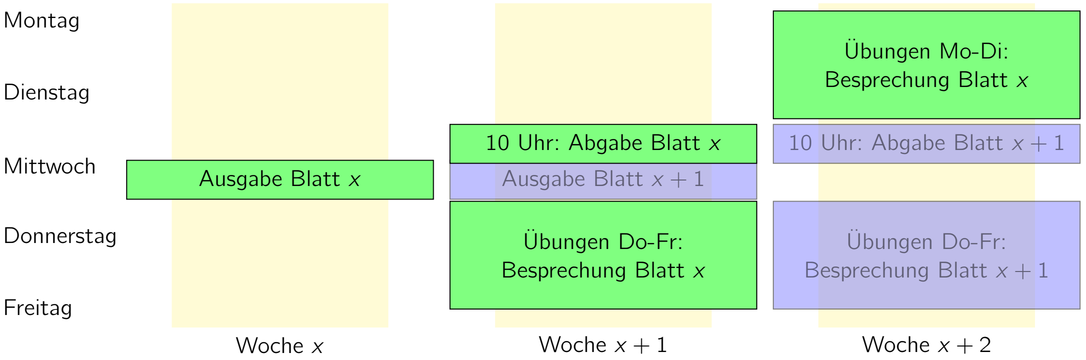
Anmeldung ist notwendig für Abgabe und Korrektur der Hausaufgaben.

Zulip-Chat

- Server-Adresse: **chat.ifi.lmu.de**
- Stream: **TCS-23S-FSK-TIMI**

Fragen und Kommentare am besten dort stellen.

Hausaufgaben



- Übungen diese Woche ab Do.: Kennenlernen+Besprechung Blatt 0 (ohne Abgabe)
- Wahl der Übungsgruppe ab heute 17:00 Uhr auf Moodle
- Abgabe und Korrektur der Übungsblätter elektronisch über Moodle
- 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 2023):

100% der erreichbaren Übungspunkte entsprechen 10% der Prüfungspunkte

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

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

Die Prüfung ist bestanden,
wenn mindestens 50% der Prüfungspunkte erreicht wurden.

Prüfungen

- Daten noch nicht bekannt
- Anmeldung zur Prüfung wird noch freigeschaltet
- Bonuspunkte gelten für Prüfung und Nachholprüfung
- Teilnahme an der Nachholprüfung auch ohne Teilnahme an der Prüfung möglich

Material

Auf der Webseite verfügbar:

- Vorlesungsfolien
- Skript zur Vorlesung:
  Markierungen mit $\star$ für nicht TIMI-relevante Teile
- Übungsblätter
- Videos von SoSe 2022

Lehrbuch:

- Uwe Schöning *Theoretische Informatik – Kurz gefasst*

Literatur

Wesentliche Quellen:

- Vorlesungsskript
- Uwe Schöning: Theoretische Informatik – kurz gefasst, 5. Auflage, Spektrum Akademischer Verlag, 2008
  Teile sind u.U. zu kurz gefasst

Weitere Literatur:

- Alexander Asteroth und Christel Baier: Theoretische Informatik, Pearson, 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, Pearson, 2006
  Der Klassiker, umfangreich (Erstauflage 1979)
- Ingo Wegener: Theoretische Informatik – eine algorithmenorientierte Einführung, 3.Auflage, Teubner Verlag, 2005.
  Gutes Buch, Algorithmen stehen im Vordergrund, Zugriff über UB

Inhaltsübersicht zur Veranstaltung

Inhalte der 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, Pumping-Lemma, Satz von Myhill und Nerode, Minimalautomaten, Abschlusseigenschaften
- Kontextfreie Sprachen: kontextfreie Grammatiken, Chomsky-Normalform, Greibach-Normalform, Pumping-Lemma, Ogdens 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 aufzählbare Sprachen
- Reduktionen
- Postsches Korrespondenzproblem

TIMI: Zum Teil nur Auswahl der Inhalte / oberflächlichere Behandlung

Inhalte: Komplexitätstheorie

- Zeitkomplexität
- Klassen P und NP
- NP-Schwere, NP-Vollständigkeit
- Polynomielle Reduktionen
- SAT-Problem
- Satz von Cook
- Weitere NP-vollständige Probleme (z.B. 3-SAT, Clique, Vertex Cover, Subset Sum, Knapsack, Directed Hqmiltonian Circuit, Hamiltonian Circuit,. . .)

TIMI: Zum Teil nur Auswahl der Inhalte / oberflächlichere Behandlung

Grundlagen: Wörter und Formale Sprachen

Wörter

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 Wörtern

- 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 der i-ten 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
  - $|abbccc| = 6$
  - $|aabbbccc| = 8$
  - $\#_a(abbccc) = 1$
  - $\#_c(aabbbccc) = 3$
- Für $w = abbbcd$ 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.

Die Konkatenation hilft folgende Mengen von Wörtern über Σ zu definieren:

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

Sei Σ ein Alphabet, dann definieren wir:

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

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

Beispiele

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

Dann ist

Beispiele

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

Dann ist

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

Beispiele

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

Dann ist

$\Sigma^0 = \{\varepsilon\}$
$\Sigma^1 = \Sigma = \{a, b\}$

Beispiele

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

Dann ist

$$\Sigma^0 = \{\varepsilon\}$$
$$\Sigma^1 = \Sigma = \{a, b\}$$
$$\Sigma^2 = \{aa, ab, ba, bb\}$$

Beispiele

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

Dann ist

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

Beispiele

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

Dann ist

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

und

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

Weitere Notationen und Begriffe

Sei w ein Wort über Σ

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

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

- $\overline{w}$ ist das rückwärts gelesene Wort w, d.h.

$$\overline{\varepsilon} = \varepsilon \text{ und für } w = a_1 \cdots a_n \text{ ist } \overline{w} = a_n \cdots a_1$$

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

Beispiele für Palindrome: anna, reliefpfeiler, lagerregal, annasusanna

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 $uw = v$.

$$\underbrace{\boxed{\quad u \quad}\boxed{\quad w \quad}}_{v}$$

- u ist ein **Suffix** von v, wenn es ein Wort w gibt mit $wu = v$.

$$\underbrace{\boxed{\quad w \quad}\boxed{\quad u \quad}}_{v}$$

- u ist ein **Teilwort** von v, wenn es Wörter w_1, w_2 gibt mit $w_1 u w_2 = v$.

$$\underbrace{\boxed{\quad w_1 \quad}\boxed{\quad u \quad}\boxed{\quad w_2 \quad}}_{v}$$

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 Σ ist eine Teilmenge von Σ^*
d.h. $L \subseteq \Sigma^*$.

Beachte: Wir verwenden L für „language“.

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: $\overline{L} := \Sigma^* \setminus L$
- Produkt: $L_1 L_2 = L_1 \circ L_2 = \{uv \mid u \in L_1 \text{ und } v \in L_2\}$

Beispiele

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

- $L_1 \cup L_2 = ?$
- $L_1 \cap L_2 = ?$
- $\overline{L_1} = ?$
- $L_1 L_2 = ?$
- $L_2 L_1 = ?$
- $L_1 L_1 = ?$

Beispiele

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

- $L_1 \cup L_2 =$ Sprache aller Wörter, die nur aus a's oder nur aus b's bestehen
- $L_1 \cap L_2 = ?$
- $\overline{L_1} = ?$
- $L_1 L_2 = ?$
- $L_2 L_1 = ?$
- $L_1 L_1 = ?$

Beispiele

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

- $L_1 \cup L_2 =$ Sprache aller Wörter, die nur aus a's oder nur aus b's bestehen
- $L_1 \cap L_2 = \{\varepsilon\}$
- $\overline{L_1} = ?$
- $L_1 L_2 = ?$
- $L_2 L_1 = ?$
- $L_1 L_1 = ?$

Beispiele

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

- $L_1 \cup L_2 =$ Sprache aller Wörter, die nur aus a's oder nur aus b's bestehen
- $L_1 \cap L_2 = \{\varepsilon\}$
- $\overline{L_1} =$ Sprache der Wörter, die mindestens ein b enthalten
- $L_1 L_2 =$?
- $L_2 L_1 =$?
- $L_1 L_1 =$?

Beispiele

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

- $L_1 \cup L_2$ = Sprache aller Wörter, die nur aus a's oder nur aus b's bestehen
- $L_1 \cap L_2 = \{\varepsilon\}$
- $\overline{L_1}$ = Sprache der Wörter, die mindestens ein b enthalten
- $L_1 L_2 = \{a^i b^j \mid i, j \in \mathbb{N}\}$
- $L_2 L_1 = ?$
- $L_1 L_1 = ?$

Beispiele

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

- $L_1 \cup L_2 =$ Sprache aller Wörter, die nur aus a's oder nur aus b's bestehen
- $L_1 \cap L_2 = \{\varepsilon\}$
- $\overline{L_1} =$ Sprache der Wörter, die mindestens ein b enthalten
- $L_1 L_2 = \{a^i b^j \mid i, j \in \mathbb{N}\}$
- $L_2 L_1 = \{b^i a^j \mid i, j \in \mathbb{N}\}$
- $L_1 L_1 = ?$

Beispiele

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

- $L_1 \cup L_2 =$ Sprache aller Wörter, die nur aus a's oder nur aus b's bestehen
- $L_1 \cap L_2 = \{\varepsilon\}$
- $\overline{L_1} =$ Sprache der Wörter, die mindestens ein b enthalten
- $L_1 L_2 = \{a^i b^j \mid i, j \in \mathbb{N}\}$
- $L_2 L_1 = \{b^i a^j \mid i, j \in \mathbb{N}\}$
- $L_1 L_1 = L_1$

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:

$$L^0 := \{\varepsilon\} \qquad L^* := \bigcup_{i \in \mathbb{N}} L^i$$

$$L^i := L \circ L^{i-1} \text{ für } i > 0 \qquad L^+ := \bigcup_{i \in \mathbb{N}_{>0}} L^i$$

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

Kleenescher Abschluss

Sei L eine Sprache. Dann ist:

$$L^0 := \{\varepsilon\} \qquad L^* := \bigcup_{i \in \mathbb{N}} L^i$$

$$L^i := L \circ L^{i-1} \text{ für } i > 0 \qquad L^+ := \bigcup_{i \in \mathbb{N}_{>0}} L^i$$

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

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

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

Kleenescher Abschluss

Sei L eine Sprache. Dann ist:

$$L^0 := \{\varepsilon\} \qquad L^* := \bigcup_{i \in \mathbb{N}} L^i$$

$$L^i := L \circ L^{i-1} \text{ für } i > 0 \qquad L^+ := \bigcup_{i \in \mathbb{N}_{>0}} L^i$$

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

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

- $L^0 = \{\varepsilon\}$
- $L^1 = L \circ L^0 = L = \{ab, ac\}$

Kleenescher Abschluss

Sei L eine Sprache. Dann ist:

$$L^0 := \{\varepsilon\} \qquad L^* := \bigcup_{i\in\mathbb{N}} L^i$$

$$L^i := L \circ L^{i-1} \text{ für } i > 0 \qquad L^+ := \bigcup_{i\in\mathbb{N}_{>0}} L^i$$

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

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

- $L^0 = \{\varepsilon\}$
- $L^1 = L \circ L^0 = L = \{ab, ac\}$
- $L^2 = \{abab, abac, acab, acac\}$

Kleenescher Abschluss

Sei L eine Sprache. Dann ist:

$$L^0 := \{\varepsilon\} \qquad L^* := \bigcup_{i \in \mathbb{N}} L^i$$

$$L^i := L \circ L^{i-1} \text{ für } i > 0 \qquad L^+ := \bigcup_{i \in \mathbb{N}_{>0}} L^i$$

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

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

- $L^0 = \{\varepsilon\}$
- $L^1 = L \circ L^0 = L = \{ab, ac\}$
- $L^2 = \{abab, abac, acab, acac\}$
- $L^3 = \{ababab, ababac, abacab, abacac, acabab, acabac, acacab, acacac\}$

Weitere Beispiele

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

Beschriebene Sprache = ?

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

Beschriebene Sprache = ?

Weitere Beispiele

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

Beschriebene Sprache = Sprache aller gültigen Uhrzeiten

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

Beschriebene Sprache = ?

Weitere Beispiele

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

Beschriebene Sprache = Sprache aller gültigen Uhrzeiten

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

Beschriebene Sprache = Sprache aller natürlichen Zahlen