PD Dr. Jan Johannsen
Elisabeth Lempa
Luca Maio

Ludwig-Maximilians-Universität München
Institut für Informatik
Besprechung 14.05.2026 bis 18.05.2026
Abgabe bis 02.06.2026, 14:00 Uhr

Übung 4 zur Vorlesung

Formale Sprachen und Komplexität

Hinweis:
Die letzte Aufgabe auf diesem Blatt ist eine Aufgabe zur Klausurvorbereitung. Diese Aufgabe orientiert sich in Form und inhaltlichen Schwerpunkten an den Klausuraufgaben. Dies bedeutet jedoch nicht, dass die anderen Aufgaben nicht klausurrelevant sind.

Die Lösungen der Klausurvorbereitungs-Aufgaben werden am Ende der Beabeitungszeit gesondert veröffentlicht, aber **nicht** im Tutorium besprochen. Die Lösungen der anderen Aufgaben werden bereits zu Beginn der Bearbeitungszeit veröffentlicht und können Ihnen bei der Bearbeitung helfen.

Wenn Sie Ihre Lösung innerhalb der Bearbeitungszeit über Moodle abgeben, erhalten Sie eine individuelle Korrektur. Die Abgabe ist freiwillig (aber nachdrücklich empfohlen).

Wenn Sie Automaten angeben, tun Sie dies immer in Form eines Zustandsgraphen. Andere Formen der Darstellung (z.B. als Liste von Übergängen) werden nicht gewertet, da sie sehr viel aufwändiger zu korrigieren sind. Vergessen Sie nicht, im Zustandsgraph Start- und Endzustände zu markieren.

Reguläre Ausdrücke sind entsprechend Definition 4.7.1 im Vorlesungsskript anzugeben.

FSK4-1 *Pumping-Lemma für reguläre Sprachen*

Zeigen Sie mit dem Pumping-Lemma für reguläre Sprachen, dass die folgenden Sprache nicht regulär ist.

$L_2 = L(G_2)$, wobei G_2 eine kontextfreie Grammatik ist mit

$$G_2 = (\{S, A, B\}, \{(,), [,]\}, P, S)$$
$$P = \{S \to (S), S \to [S], S \to A, S \to B, A \to (), A \to [], B \to S, B \to BB\}$$

L_2 ist die Sprache der zueinander passenden eckigen und runden Klammern, d.h. es sind z.B. $([])[] \in L_2$ und $()() \in L_2$, aber $([\notin L_2$ und $) \notin L_2$ (vgl. Aufgabe FSK1-3).

FSK4-2 ***Reguläre Ausdrücke und Abschlusseigenschaften***

a) Betrachten Sie den regulären Ausdruck $\alpha = (a|b)^*(ab|ba)(a|b)^*$.

   i) Geben Sie einen NFA ohne ε-Übergänge an, der $L(\alpha)$ erkennt. Sie können die Algorithmen aus der Vorlesung zur Konstruktion eines NFA aus einem regulären Ausdruck und zur Elimination von ε-Übergängen verwenden, müssen aber nicht.

   ii) Geben Sie einen DFA an, der $L(\alpha)$ erkennt. Sie können die Potenzmengenkonstruktion verwenden, müssen aber nicht.

b) Geben Sie einen regulären Ausdrucke an, der die Sprache L_3 erkennt, also die Sprache der Wörter über dem Alphabet $\Sigma_1 = \{a, b, c\}$, die mit a oder b anfangen und mindestens ein c enthalten.

c) Zeigen Sie mithilfe der Abschlusseigenschaften regulärer Sprachen, dass die Sprache $L_5 = \{a^i w d^{i+1} \mid i \in \mathbb{N},\ w \in \{b, c\}^*\}$ über dem Alphabet $\Sigma_3 = \{a, b, c, d\}$ nicht regulär ist. Sie dürfen annehmen, dass die Sprache $L_1 = \{a^i b^j c^k d^i \mid i, j, k \in \mathbb{N}_{>0}\}$ aus Aufgabe FSK4-Kb) nicht regulär ist.

FSK4-3 ***Grammatik über Automaten zu Grammatik***

Gegeben sei die reguläre Grammatik

$$G = (\{S, A, B, C\}, \{a, b\}, \{S \to aA \mid bB,\ A \to bB,\ B \to bC,\ C \to aC \mid a\}, S)$$

a) Erzeugen Sie gemäß der Konstruktion aus der Vorlesung aus G einen NFA A mit $L(G) = L(A)$.

b) Erzeugen Sie mit der Potenzmengenkonstruktion aus A einen DFA B mit $L(B) = L(A)$. Geben Sie nur den vom Startzustand erreichbaren Teil von A an.

c) Erzeugen Sie gemäß der Konstruktion aus der Vorlesung aus B eine Grammatik H mit $L(B) = L(H)$.

d) Vergleichen Sie die Grammatiken G und H. Beschreiben Sie die Gemeinsamkeiten dieser Grammatiken, sowie ihre Unterschiede.

   Überlegen Sie sich, wodurch diese Effekte zustande kommen.

FSK4-4 ***DNA-Analyse mit NFA***

Diese Aufgabe handelt von der Analyse von Desoxyribonukleinsäure (DNS/DNA) mithilfe von NFA. DNA ist eine Abfolge der Basen Adenin, Thymin, Guanin und Cytosin, typischerweise mit A, T, G und C abgekürzt. Dementsprechend ist das Alphabet aller Automaten in dieser Aufgabe $\Sigma = \{A, C, G, T\}$.

a) Um das Vorkommen einer Basensequenz zu finden, wird aus dieser Sequenz ein NFA erzeugt, der alle Wörter akzeptiert, in denen diese Sequenz als Teilwort vorkommt.

   Geben Sie einen NFA B an, der genau diejenigen Wörter akzeptiert, in denen $ACTC$ als Teilwort vorkommt.

   Hinweis: Sie können (müsssen aber nicht) dazu den regulären Ausdruck $(A|C|G|T)^*ACTC(A|C|G|T)^*$ verwenden, der genau diese Sprache akzeptiert.

   Hinweis: Sie können auch einen DFA angeben, aber ein NFA ist übersichtlicher.

b) Beim Kopieren von DNA kann es vorkommen, dass Fehler auftreten. Zum Beispiel kann eine Base durch eine andere ersetzt werden; es kann eine Base ausgelassen werden; es kann eine zusätzliche Base eingefügt werden; und es können auch komplexere Fehler auftreten. Zur Vereinfachung behandeln wir hier nur den Fall, dass eine Base durch eine andere ersetzt wird.

   Aus einem NFA $D = (Z_D, \Sigma, \delta_D, S_D, E_D)$ kann ein NFA $F = (Z_F, \Sigma, \delta_F, S_F, E_F)$ erzeugt werden, der alle Wörter akzeptiert, die durch höchstens k fehlerhafte Ersetzungen aus D entsteht.

   Dabei sind

   - $Z_F = Z_D \times \{0, \ldots, k\}$
   - $\delta_F((q,i),a) = \quad \{(q',i) \mid q' \in \delta_D(q,a)\}$
     $\qquad\qquad \cup \{(q',i+1) \mid \exists b \in \Sigma.\, q' \in \delta_D(q,b) \wedge i+1 \leq k\}$
   - $S_F = S_D \times \{0\} = \{(s,0) \mid s \in S_D\}$
   - $E_F = E_D \times \{0, \ldots, k\}$

   Berechnen Sie mit der obigen Konstruktion einen NFA H aus B, der Wörter mit bis zu 2 Fehlern akzeptiert.

c) Geben Sie an und begründen Sie, welche der folgenden Wörter von H akzeptiert werden. Prüfen Sie, ob Ihr Ergebnis korrekt ist, also ob die erkannten Wörter tatsächlich diejenigen sind, bei denen bis auf höchstens 2 Fehler das Wort $ACTC$ als Teilwort vorkommt.

   $AAAACCCAAA$, $GAGGCGT$, $TAGCA$, $TCTCA$

d) Beweisen Sie, dass die Konstruktion aus b) korrekt ist, also tatsächlich für jeden NFA D und jedes k einen NFA F liefert, der maximal k Fehler zulässt.

   Hinweis: Sie können auch als Vorüberlegung dies erst für den NFA H zeigen.

Klausurvorbereitung FSK-4-K

a) Geben Sie einen regulären Ausdruck an, der die Sprache L_4 erkennt. L_4 ist die Sprache der Wörter über dem Alphabet $\Sigma_2 = \{a, b\}$, die keine zwei a's hintereinander enthalten.

b) Zeigen Sie mit dem Pumping-Lemma für reguläre Sprachen, dass die Sprache $L_1 = \{a^i b^j c^k d^i \mid i, j, k \in \mathbb{N}_{>0}\}$ über dem Alphabet $\Sigma_1 = \{a, b, c, d\}$ nicht regulär ist.