Prof. Dr. Jasmin Blanchette
Elisabeth Lempa
Luca Maio

Ludwig-Maximilians-Universität München
Institut für Informatik
Besprechung 30.05.2025 bis 02.06.2025
Abgabe bis 09.06.2025, 10:00 Uhr

Lösungsvorschlag zur Übung 5 zur Vorlesung
Theoretische Informatik für Studierende der Medieninformatik

TIMI5-1 *Reguläre und nicht-reguläre Sprachen*

a) Zeigen Sie, dass die Sprache $\{a^i b^j c^k \mid i, j, k \in \mathbb{N} \text{ und wenn } i = 2, \text{ dann } j < k\}$ über dem Alphabet $\Sigma = \{a, b, c\}$ die Pumping-Eigenschaft erfüllt.

LÖSUNGSVORSCHLAG:
Wir wählen $n = 4$.
Sei $z = a^i b^j c^k \in L$ mit $|z| \geq n$.
Wir müssen für jedes solche z eine Zerlegung $z = uvw$ angeben mit $|uv| \leq n$, $|v| \geq 1$ und $uv^l w \in L$ für alle $l \in \mathbb{N}$.
Vollständige Fallunterscheidung:

- $i = 2$. Wähle die Zerlegung $u = \varepsilon$, $v = aa$, $w = b^j c^k$. Es ist $uv^l w \in L$ für jedes $l \in \mathbb{N}$, denn:
  - Für $l = 1$ ist $uv^l w = z \in L$.
  - Für $l \neq 1$ enthält $uv^l w$ nicht genau 2 a's und somit ist die Anzahl der b's und c's wieder irrelevant.
- $i \geq 4$. Wähle die Zerlegung $u = a^3$, $v = a$, $w = a^{i-4} b^j c^k$ mit $|uv| \leq 4$ und $|v| \geq 1$. Es ist $uv^l w \in L$ für jedes $l \in \mathbb{N}$, da die Anzahl der a's in $uv^l w$ immer mindestens 3 ist und somit die Anzahl der b's und c's irrelevant ist.
- $i < 4$ und $i \neq 2$. Wegen $|z| \geq 4$ ist $j \geq 1$ oder $k \geq 1$.
  - Wenn $j \geq 1$: Wähle die Zerlegung $u = a^i$, $v = b$, $w = b^{j-1} c^k$ mit $|uv| \leq 4$ und $|v| \geq 1$. Es ist $uv^l w \in L$ für jedes $l \in \mathbb{N}$, da die Anzahl der a's in $uv^l w$ immer $i \neq 2$ bleibt und somit die Anzahl der b's und c's irrelevant ist.
  - Wenn $j = 0$: Wähle die Zerlegung $u = a^i$, $v = c$, $w = c^{k-1}$. Diese Zerlegung erfüllt mit analoger Begründung alle Bedingungen.

b) Sind die folgenden Sprachen L_i, $i \in \{1, 2, 3\}$, über den Alphabeten Σ_i regulär? Wenn ja, geben Sie einen regulären Ausdruck an, der L_i erkennt. (Sie müssen nicht beweisen, dass der reguläre Ausdruck L_i erkennt.) Wenn nein, zeigen Sie die Nichtregularität mit dem Pumping-Lemma für reguläre Sprachen.

i) $L_1 = \{ac^iba^jb \mid i, j \in \mathbb{N}\}$ mit $\Sigma_1 = \{a, b, c\}$

> **LÖSUNGSVORSCHLAG:**
> Regulär; $L_1 = L(ac^*ba^*b)$.

ii) $L_2 = \{a^pb^p \mid p \in \mathbb{N} \text{ ist prim}\}$ mit $\Sigma_2 = \{a, b\}$

> **LÖSUNGSVORSCHLAG:**
> Nicht regulär. Beweis mit dem Pumping-Lemma.
> Sei $n \in \mathbb{N}_{>0}$ beliebig.
> Wir wählen $z \in L_2$ als $z = a^pb^p$, wobei p die kleinste Primzahl mit $p \geq n$ ist. Damit ist $|z| \geq n$.
> Sei $z = uvw$ eine beliebige Zerlegung von z, sodass $|uv| \leq n$, $|v| \geq 1$ und $uv^iw \in L_2$ für jedes $i \in \mathbb{N}$. Da $|uv| \leq n$ ist, ist $v = a^k$ für ein $k \in \mathbb{N}_{>0}$.
> Wir wählen $i = 0$. Das Wort uv^iw hat weniger a's als b's und somit ist $uv^iw \notin L_2$. Widerspruch.

iii) $L_3 = \{a^{2n+1} \mid n \in \mathbb{N}\}$ mit $\Sigma_3 = \{a\}$.

> **LÖSUNGSVORSCHLAG:**
> Regulär; $L_3 = L(a(aa)^*)$.

TIMI5-2 ***Konservative Erweiterungen regulärer Ausdrücke***

In der Praxis werden reguläre Ausdrücke häufig mit weiteren Operatoren erweitert. Eine solche Erweiterung ist *konservativ*, wenn die erweiterten regulären Ausdrücke nur reguläre Sprachen beschreiben. Geben Sie in jeder Teilaufgabe an, ob die beschriebene Erweiterung konservativ ist, und beweisen Sie Ihre Antwort. Dabei sei α ein regulärer Ausdruck über einem beliebigen Alphabet.

a) $\alpha?$: Teilwörter, die von α erkannt werden, dürfen vorkommen, müssen aber nicht. Die Semantik von $\alpha?$ ist also $L(\alpha?) = \{\varepsilon\} \cup L(\alpha)$.

> **LÖSUNGSVORSCHLAG:**
> Die Erweiterung ist konservativ. Es ist $L(\alpha?) = L(\alpha|\varepsilon)$, das heißt jeder Teilausdruck $\alpha?$ kann durch den regulären Ausdruck $\alpha|\varepsilon$ ersetzt werden, ohne die Bedeutung des gesamten Ausdrucks zu ändern.

b) α^+: wie α^*, aber α muss mindestens einmal vorkommen.

$$L(\alpha^+) = \bigcup_{i \in \mathbb{N}_{>0}} L(\alpha)^i = L(\alpha) \cup L(\alpha)^2 \cup L(\alpha)^3 \cup \cdots$$

LÖSUNGSVORSCHLAG:
Konservativ, denn $L(\alpha^+) = L(\alpha\alpha^*)$.

c) $\alpha^{\{i,j\}}$ mit $i, j \in \mathbb{N}$ und $i \leq j$: wie α^*, aber α muss mindestens i-mal und darf höchstens j-mal wiederholt werden.

$$L(\alpha^{\{i,j\}}) = \bigcup_{k=i}^{j} L(\alpha)^k = L(\alpha)^i \cup L(\alpha)^{i+1} \cup L(\alpha)^{i+2} \cup \cdots \cup L(\alpha)^j$$

LÖSUNGSVORSCHLAG:
Konservativ, denn

$$L(\alpha^{\{i,j\}}) = L(\underbrace{\alpha\alpha\ldots\alpha}_{i\text{-mal}}\underbrace{\alpha?\alpha?\ldots\alpha?}_{(j-i)\text{-mal}})$$

d) $\backslash n$ mit $n \in \mathbb{N}$. In einem regulären Ausdruck α bezeichnen wir den n-ten Teilausdruck der Form (α_0) (wobei α_0 ein regulärer Ausdruck ist) als die n-te Capturing Group. Ein Teilausdruck $\backslash n$ in α wird dann als Backreference bezeichnet und erkennt genau die Zeichenkette, die von α_0 erkannt wurde. Beispielsweise erkennt $(a|b)\backslash 1$ die Wörter aa und bb, aber nicht ab oder ba.

LÖSUNGSVORSCHLAG:
Nicht konservativ. Der Ausdruck $((a|b)^*)\backslash 1$ erkennt die Sprache $\{ww \mid w \in \{a,b\}^*\}$, die bekanntlich nicht regulär ist.