Prof. Dr. Jasmin Blanchette
Jannis Limperg
Luca Maio

Ludwig-Maximilians-Universität München
Institut für Informatik
Abgabe bis 29. Mai 2024, 10:00 Uhr

Übung 5 zur Vorlesung

Formale Sprachen und Komplexität

FSK5-1 ***Myhill und Nerode*** (2 Punkte)

a) Sei $L = L(ab^*c)$ eine Sprache über dem Alphabet $\Sigma = \{a, b, c\}$. Geben Sie für jedes der folgenden Wörter u_i, $i \in \{1, 2, 3\}$, eine kompakte Beschreibung der Äquivalenzklasse $[u_i]$ der Nerode-Relation von L an.

i) $u_1 = abb$
ii) $u_2 = \varepsilon$
iii) $u_3 = c$

b) Bestimmen Sie den Nerode-Index folgender Sprachen $L_i, i \in \{1, 2, 3\}$, über Alphabeten Σ_i und entscheiden Sie mit dem Satz von Myhill und Nerode, welche der Sprachen regulär sind. Geben Sie für jede Sprache mit endlichem Nerode-Index alle paarweise verschiedenen Äquivalenzklassen an (1 Repräsentant pro Klasse).

i) $L_1 = \{aaab, aabb, abab, abbb, baab, babb, bbab, bbbb\}$ mit $\Sigma_1 = \{a, b\}$
ii) $L_2 = \{w\overline{w}w \mid w \in \Sigma_2^*\}$ mit $\Sigma_2 = \{a, b\}$
iii) $L_3 = \{a^i b^j c^k \mid i, j, k \in \mathbb{N}$ und wenn $i = 2$, dann $j < k\}$ mit $\Sigma_3 = \{a, b, c\}$

FSK5-2 ***Pumping-Lemma für kontextfreie Sprachen*** (2 Punkte)

Zeigen Sie mit dem Pumping-Lemma für kontextfreie Sprachen, dass die Sprache $L = \{w\overline{w}w \mid w \in \Sigma^*\}$ über dem Alphabet $\Sigma = \{a, b\}$ nicht kontextfrei ist.

FSK5-3 ***Reguläre und nicht-reguläre Sprachen*** (0 Punkte)

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

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) $L_1 = \{ac^i ba^j b \mid i, j \in \mathbb{N}\}$ mit $\Sigma_1 = \{a, b, c\}$
ii) $L_2 = \{a^p b^p \mid p \in \mathbb{N}$ ist prim$\}$ mit $\Sigma_2 = \{a, b\}$
iii) $L_3 = \{a^{2n+1} \mid n \in \mathbb{N}\}$ mit $\Sigma_3 = \{a\}$.

FSK5-4 ***Konservative Erweiterungen regulärer Ausdrücke*** (0 Punkte)

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 $L(\alpha?) = \{\varepsilon\} \cup L(\alpha)$.

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

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

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.

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

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 ab, aber nicht ab oder ba.