Technische Universität München
Fakultät für Informatik
Lehrstuhl Informatik XIV
Prof. Dr. Ernst W. Mayr
Dr. Werner Meixner

Sommersemester 2015
Übungsblatt 5
11. Mai 2015

Theoretische Informatik

Abgabetermin: 18. Mai 2015, 13 Uhr in die **THEO Briefkästen**

Hausaufgabe 1 (4 Punkte)

Sei $\Sigma = \{0, 1, (,), +, ^*, \boldsymbol{\emptyset}, \boldsymbol{\epsilon}\}$ die Zeichenmenge, aus der reguläre Ausdrücke über dem Alphabet $\{0, 1\}$ gebildet werden. Wir schreiben hier $+$ anstelle von $|$, um Zeichenverwirrungen zu vermeiden.

Die induktive Definition der regulären Ausdrücke kann man als eine Grammatik $G = (\{R\}, \Sigma, P, R)$ auffassen, deren Produktionenmenge P aus den folgenden Produktionen besteht:

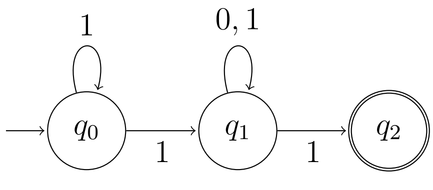
$$R \to 0 \mid 1 \mid \boldsymbol{\epsilon} \mid \boldsymbol{\emptyset} \mid RR \mid R + R \mid R^* \mid (R)$$

1. Zeigen Sie, dass die Grammatik G nicht eindeutig ist.
2. Geben Sie eine eindeutige Grammatik G' für die regulären Ausdrücke an, die die Bindungsstärken in regulären Ausdrücken respektiert (also z.B. Konkatenation stärker als $+$ bindet).

   Begründen Sie, dass Ihre Grammatik eindeutig ist.
3. Geben Sie den Syntaxbaum für das Wort 01^*0+1 bezüglich Ihrer eindeutigen Grammatik an.

Hausaufgabe 2 (4 Punkte)

Durch die folgende Grafik sei ein nichtdeterministischer endlicher Automat $A = (Q, \Sigma, \delta, \{q_0\}, F)$ gegeben.



1. Bestimmen Sie $\hat{\delta}(\{q_0, q_1, q_2\}, 0)$! Geben Sie einen regulären Ausdruck α an, der die von A akzeptierte Sprache beschreibt, d. h. $L(\alpha) = L(A)$.
2. Konstruieren Sie mit dem Potenzmengenverfahren einen deterministischen endlichen Automaten B, der $L(A)$ akzeptiert.

   Stellen Sie den erhaltenen Automaten B als Übergangsgraph dar.
3. Wie ist der Automat B abzuändern, so dass B das Komplement von $L(A)$ akzeptiert, d. h., so dass $L(B) = \{0, 1\}^* \setminus L(A)$ gilt?

Hausaufgabe 3 (4 Punkte)

Sei Σ eine nichtleere Zeichenmenge.

1. Sei L die Menge aller Wörter über Σ mit geradzahliger Länge. Geben Sie ein Verfahren an, das für einen beliebigen DFA $A = (Q, \Sigma, \delta, q_0, F)$ entscheidet, ob jedes (oder nicht jedes) von A akzeptierte Wort eine ungerade Länge besitzt.

2. Sei $E(x)$ ein Eigenschaft für Wörter x über Σ, so dass $K = \{x\,;\, E(x)\}$ eine reguläre Sprache ist. Verallgemeinern Sie Ihr obiges Verfahren so, dass es für einen beliebigen DFA $A = (Q, \Sigma, \delta, q_0, F)$ entscheidet, ob jedes (oder nicht jedes) von A akzeptierte Wort x die Eigenschaft E besitzt.

Hausaufgabe 4 (4 Punkte)

Stellen Sie die folgenden Sprachen jeweils durch einen regulären Ausdruck dar:

1. Die Sprache L_1 aller Wörter über $\{0, 1\}$, die mit 101 beginnen und mit 010 enden.

2. Die Sprache $L_2 = \{a^m b^n\,;\, (m + n) \bmod 3 = 1\}$.

3. Die Sprache L_3 aller Wörter über $\{0, 1\}$, in denen kein Paar aufeinanderfolgender Nullen weiter links steht als ein beliebiges Paar benachbarter Einsen.

   Hinweis: Sei $w \in L_3$ mit $w = x_1 x_2 \ldots x_n$ und $x_i \in \{0, 1\}$. Sei i der größte Index (größte Zahl), so dass $x_i x_{i+1} = 11$. i gibt also die Position des am weitesten rechts stehenden Einserpaares. Falls kein Einserpaar existiert, dann setzen wir $i = 0$.
   Dann ist klar, dass das Wort $v = x_{i+1} x_{i+2} \ldots x_n$ kein Einserpaar enthält. Andererseits darf nach Definition von L_3 das Wort $u = x_1 x_2 \ldots x_i$ kein Paar von Nullen enthalten.

Hausaufgabe 5 (4 Punkte)

Das Reißverschlussprodukt von zwei Sprachen L_1 und L_2 ist wie folgt definiert:

$$L_1 \,\%\, L_2 = \{a_1 b_1 a_2 b_2 \ldots a_n b_n\,;\, a_i \in \Sigma, b_i \in \Sigma, a_1 \ldots a_n \in L_1, b_1 \ldots b_n \in L_2\}$$

1. Geben Sie einen regulären Ausdruck an, der $L(aba^*) \,\%\, L(bc^*)$ beschreibt.

2. Zeigen Sie durch eine geeignete Automatenkonstruktion: Wenn L_1 und L_2 regulär sind, dann ist auch $L_1 \,\%\, L_2$ regulär.

   Beschreiben Sie Ihre Konstruktion zunächst informell. Für die Konstruktion bildet man das Produkt der beiden Automaten mit einem zusätzlichen Flag, welches angibt, welcher der beiden Automaten gerade „an der Reihe“ ist.

Zusatzaufgabe 3 (wird nicht korrigiert)

Das sogenannte Shuffle-Produkt spielt in der Theorie der nebenläufigen Systeme eine wichtige Rolle. Für zwei Sprachen L_1 und L_2 bezeichnet $L_1 \parallel L_2$ die Menge der Wörter, die man erhält, indem man zwei Wörter $v \in L_1$ und $w \in L_2$ beliebig miteinander verschränkt. Dabei können sich Teile aus v und w beliebig abwechseln, wobei die Reihenfolge der Zeichen aus v und w jedoch erhalten bleibt. Das kann man sich gut als das Ineinanderschieben zweier Kartenstapel veranschaulichen. Formal definieren wir $L_1 \parallel L_2$ wie folgt:

$$L_1 \parallel L_2 = \{v_1 w_1 v_2 w_2 \cdots v_n w_n \,;\ v_1, \ldots, v_n, w_1, \ldots, w_n \in \Sigma^*, \\ v_1 v_2 \cdots v_n \in L_1 \text{ und } w_1 w_2 \cdots w_n \in L_2\}$$

1. Versuchen Sie, eine einfache Beschreibung von $L((01)^*) \parallel L((10)^*)$ zu finden.
2. Begründen Sie: Wenn L_1 und L_2 regulär sind, dann ist auch $L_1 \| L_2$ regulär.
   Hinweis: Konstruieren Sie einen NFA für $L_1 \parallel L_2$.
3. Führen Sie die Konstruktion konkret für die Sprachen $L((01)^*)$ und $L((10)^*)$ durch und geben Sie einen regulären Ausdruck an, der die Sprache beschreibt.
4. Vergleichen Sie dass Reißverschlussprodukt mit dem Shuffle-Produkt und finden Sie zwei Sprachen A und B und ein Wort, das in $(A \parallel B)$, aber nicht in $A \,\%\, B$ liegt.

Hinweis: *Die Vorbereitungsaufgaben bereiten die Tutoraufgaben vor und werden in der Zentralübung unterstützt. Tutoraufgaben werden in den Übungsgruppen bearbeitet. Hausaufgaben sollen selbstständig bearbeitet und zur Korrektur und Bewertung abgegeben werden.*

Vorbereitung 1

1. Gegeben seien $\Sigma = \{d, e, f\}$ und $L = \{d^k e^m f^n \;;\; k, m, n \in \mathbb{N}, m \neq n\}$.
   Zeigen Sie, dass die Sprache L kontextfrei ist.
2. Gegeben seien $\Sigma = \{d, e, f\}$ und $L = \{d^m e^m f^n \;;\; m, n \in \mathbb{N}, m \neq n\}$.
   Zeigen Sie, dass die Sprache L ein Durchschnitt kontextfreier Sprachen ist.

Vorbereitung 2

(Wiederholungsaufgabe)
Wandeln Sie die durch folgende Produktionen gegebene Grammatik mit Startsymbol S in Chomsky-Normalform um:

$$\begin{array}{lllllll} S & \to & 0A0 \mid 1B1 \mid BB\,, & & & & \\ A & \to & D\,, & B & \to & S \mid A\,, \\ C & \to & S \mid \epsilon\,, & D & \to & C\,. \end{array}$$

Tutoraufgabe 1 (CYK)

Wir betrachten die kontextfreie Grammatik $G = (V, \{a, b\}, P, S)$ mit Produktionen in Chomsky-Normalform

$$\begin{aligned}
S &\rightarrow YC_3 \mid YX \mid AC_3 \mid AX \mid YY \mid YB \mid YA \mid AY \mid AB \mid AA\,,\\
X &\rightarrow ZC_2 \mid ZB \mid BZ \mid BB\,,\\
Y &\rightarrow AC_1 \mid AY \mid AA\,,\\
C_1 &\rightarrow YZ \mid AZ\,,\\
C_2 &\rightarrow BZ\,,\\
Z &\rightarrow YC_3 \mid YX \mid AC_3 \mid AX \mid YY \mid YB \mid YA \mid AY \mid AB \mid AA\,,\\
C_3 &\rightarrow XZ \mid YZ \mid BZ \mid AZ\,,\\
A &\rightarrow a\,,\\
B &\rightarrow b\,.
\end{aligned}$$

Entscheiden Sie mit Hilfe des Cocke-Younger-Kasami-Algorithmus, ob die Wörter *aabaa* und *abab* in der Sprache $L(G)$ enthalten sind! Geben Sie gegebenenfalls Ableitungen an!

Tutoraufgabe 2 (Produktive Regeln)

Sei G eine kontextfreie Grammatik in CNF, die nur nützliche Symbole enthält. Man zeige: $L(G)$ ist genau dann unendlich, wenn G einen Zyklus enthält, d. h. ein Nichtterminal A, so dass $A \underset{G}{\rightarrow}{}^{n} \alpha A \beta$ mit $n > 0$ gilt.

Tutoraufgabe 3 (Pumping-Lemma für CFL)

1. Zeigen Sie, dass $L = \{a^n b^m c^k\,;\, 0 < n < m < k\}$ keine kontextfreie Sprache ist.
2. Sei $f : \mathbb{N} \rightarrow \mathbb{N}$ eine beliebige Funktion und es sei $L = \{a^n b^{f(n)} c^n\,;\, n \in \mathbb{N}\}$.

   Beweisen Sie:

   Falls L kontextfrei ist, dann ist f beschränkt, d. h., dann gilt

   $$(\exists k \in \mathbb{N}\ \forall n \in \mathbb{N})\,[f(n) \leq k]\,.$$