Prof. Dr. Hans-Jörg Kreowski, Sabine Kuske
Studiengang Informatik
MZH 3260, 3180 Tel.: 2956, 2335, 3697 (Sekr.), Fax: 4322
email:{kreo,kuske}@informatik.uni-bremen.de

Dezember 1999
Aufgaben 4

Theoretische Informatik II

4. Übungsblatt

Reguläre Sprachen über einem Alphabet I lassen sich syntaktisch mittels regulärer Ausdrücke beschreiben. Die Menge $REG(I)$ aller regulären Ausdrücke über I ist dabei rekursiv wie folgt definiert:

i. $\mathit{empty} \in REG(I)$

ii. $\mathit{lambda} \in REG(I)$

iii. $x \in I$ impliziert $x \in REG(I)$

iv. $\mathit{expr}_1, \mathit{expr}_2 \in REG(I)$ impliziert $(\mathit{expr}_1 \; + \; \mathit{expr}_2) \in REG(I)$

v. $\mathit{expr}_1, \mathit{expr}_2 \in REG(I)$ impliziert $(\mathit{expr}_1 \; \circ \; \mathit{expr}_2) \in REG(I)$

vi. $\mathit{expr} \in REG(I)$ impliziert $(\mathit{expr}^*) \in REG(I)$.

Die Ausdrücke *empty* und *lambda* sind Zeichen, von denen keines in I liegt. Ebenso liegt keines der Sonderzeichen (,), $+$, $\circ$, * in I. Beachte, daß die Definition so gewählt ist, daß jeder Ausdruck eindeutig rekursiv aufgebaut ist. Entweder handelt es sich um genau einen der atomaren Ausdrücke (i – iii), was nach Wahl der Zeichen eindeutig ist. Oder der Ausdruck ist zusammengesetzt (iv – vi), was sich eindeutig rekonstruieren läßt (z.B. durch Induktion über die Anzahl der in den Ausdrücken enthaltenen Operatoren). Also sowohl welcher Fall eintritt als auch welche Ausdrücke eingebaut sind, läßt sich genau ermitteln. Deshalb ist jede Zuordnung, die in Abhängigkeit vom rekursiven Aufbau definiert wird, eine Abbildung, wie beispielsweise die folgende.

Die durch einen regulären Ausdruck $\mathit{expr} \in REG(I)$ beschriebene reguläre Sprache über I wird mit $L(\mathit{expr})$ bezeichnet, wobei

i. $L(\mathit{empty}) = \emptyset$

ii. $L(\mathit{lambda}) = \{\lambda\}$

iii. $L(x) = \{x\}$

iv. $L((\mathit{expr}_1 \; + \; \mathit{expr}_2)) = L(\mathit{expr}_1) \cup L(\mathit{expr}_2)$

v. $L((\mathit{expr}_1 \; \circ \; \mathit{expr}_2)) = L(\mathit{expr}_1)L(\mathit{expr}_2)$

vi. $L((\mathit{expr}^*)) = L(\mathit{expr})^*$.

1. (a) Konstruiere eine kontextfreie Chomsky-Grammatik

$$G_{REG} = (N_{REG}, T_{REG}, P_{REG}, <\mathit{reg}>),$$

die genau alle regulären Ausdrücke über I erzeugt. Wähle dabei

$$\begin{aligned} N_{REG} &= \{<\mathit{reg}>\} \text{ und} \\ T_{REG} &= I \cup \{\mathit{empty}, \mathit{lambda}, (,), +, \circ, ^*\}. \end{aligned}$$

(20%)

(b) Sowohl die oben angegebene rekursive Definition für reguläre Ausdrücke als auch G_{REG} liefern ein Verfahren zur Konstruktion dieser Ausdrücke. Worin unterscheiden sich die beiden Konstruktionsverfahren? (10%)

2. Schreibe einen regulären Ausdruck über $\{a, b\}$, der die Sprache

$$\{a, b\}^* - \{a^m b^n \mid m, n \geq 1\}$$

beschreibt, und erläutere den Lösungsweg. (Hinweis: Um Klammern weglassen zu können, darf angenommen werden, daß $*$ stärker bindet als $\circ$, welches wiederum stärker bindet als $+$. (30%)

3. Konstruiere Kellerautomaten, welche die folgenden Sprachen erkennen:

   (a) $\{w\$\mathit{trans}(w) \mid w \in \{a, b\}^*\}$,

   (b) $\{w\mathit{trans}(w) \mid w \in \{a, b\}^*\}$,

   (c) $\{w \in \{a, b\}^* \mid \mathit{count}(a, w) = 2 \cdot \mathit{count}(b, w)\}$,

   Gib für den Kellerautomaten aus (a) eine Konfigurationsfolge mit dem Eingabewort $aabb\$bbaa$ an. (50%)

Die bearbeiteten Übungsaufgaben sind spätestens am 20.1.00 abzugeben.