Übung zur Vorlesung

Theoretische Informatik I

Prof. Dr. Christoph Kreitz
Universität Potsdam, Theoretische Informatik, WS 2008/09

Blatt 4 (Version 2) — Abgabetermin: 24.11.2008, 12:30 Uhr

Vorbereitung auf die nächste Vorlesung: Arbeiten Sie sich in das Thema "Grammatiken" ein. Verwenden Sie hierzu z.B. die Folien der Einheit 2.4, eines der empfohlenen Bücher oder das Internet.

Aufgabe 4.1 (Elemente der Sprachen regulärer Ausdrücke)

Geben Sie für jeden der folgenden regulären Ausdrücke über $\Sigma = \{a, b, c\}$ je zwei Wörter an, die in der durch ihn repräsentierten Sprache enthalten sind sowie zwei Wörter, die nicht enthalten sind.

1. $(a + b)^*$
2. $a(b)^*$
3. $(ab)^*$
4. $(a + b + c)^*$
5. $(a + \varepsilon)b^*$
6. a^*b^*.
7. $a(ab)^*b$.
8. $(a + ba + bb)(a + b)^*$.

Aufgabe 4.2 (Analyse regulärer Ausdrücke)

Gegeben sei der reguläre Ausdruck $R = 000^*(1+\varepsilon)0^* + 00^*(1+\varepsilon)00^* + 0^*(1+\varepsilon)000^*$. Beschreiben Sie die Sprache $L_R = L(R)$ und zeigen Sie, dass L_R tatsächlich die Sprache des Ausdrucks R ist!

Aufgabe 4.3 (Konstruktion regulärer Ausdrücke)

Geben Sie reguläre Ausdrücke für die folgenden Sprachen an.

1. $\{w \in \{0, 1\}^* \mid$ die Anzahl der Einsen in w ist durch 3 teilbar$\}$
2. $\{w \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^* \mid w$ enthält die Symbole $\mathtt{a}$ oder $\mathtt{c}\}$
3. $\{w \in \{0, 1\}^* \mid w$ beginnt mit 0 und $|w|$ ist gerade oder w beginnt mit 1 und $|w|$ ist ungerade$\}$

Geben Sie für mindestens eine dieser Sprachen einen endlichen Automaten an, der sie akzeptiert.

Aufgabe 4.4 (Äquivalenz regulärer Ausdrücke)

Zeigen oder widerlegen Sie die Gültigkeit der folgenden Äquivalenzen (für beliebige reguläre Ausdrücke R und S):

1. $(R + S) + R \doteq S + R$
2. $(R + S)^* \doteq R^* + S^*$
3. $\varepsilon + (R + S)^+ \doteq ((R + S)(R + S)^*)^*$

Aufgabe 4.5 (Etwas zum Nachdenken: Grenzen endlicher Automaten)

Die Berechnungsstärke endlicher Automaten wird durch die Endlichkeit ihrer Zustandsmenge beschränkt. Versuchen Sie, eine Sprache zu konstruieren, die von keinem endlichen Automaten akzeptiert werden kann und geben Sie eine informale Begründung dafür an.

Hausaufgabe 4.6 (Konstruktion regulärer Ausdrücke für praktische Anwendungen)

Sie wollen im Internet nach preisgünstigen Musterlösung zu den Hausaufgaben für "Theoretische Informatik I" suchen. Dazu müssen Sie einen regulären Ausdruck konstruieren, der Ihnen passende Textmuster erkennt. Ein solches Textmuster wäre zum Beispiel

TI-1 Lösungen zu den Hausaufgaben (außer Aufgabe 4.6) für nur 23.50 €

Es reicht aus Textmuster zu erkennen, die mit "TI-1" beginnen und mit einem Preis enden. Preise können z.B. die folgende Form haben: `5 €`, `0.99 €`, `9.- €`, `€ 7`, `€ 1.-`, `€ 19.99`.

Beschreiben Sie die Gestalt der möglichen Textmuster informal und geben Sie dann einen regulären Ausdruck an, der diese Textmuster beschreibt. Erklären Sie dabei, welche Bestandteile der informaeln Beschreibung durch welchen Teilausdruck beschrieben werden. Für den regulären Ausdruck (0+1+2+3+4+5+6+7+8+9) können Sie die abkürzende Schreibweise [0-9] verwenden. Entsprechend können Sie Groß- bzw. Kleinbuchstaben (inklusive der Sonderzeichen) mit [A-Z] bzw. [a-z] abkürzen.

Hausaufgabe 4.7 (Äquivalenz regulärer Ausdrücke)

Zeigen oder widerlegen Sie die Gültigkeit der folgenden Äquivalenzen (für beliebige reguläre Ausdrücke R und S):

1. $(R+S)^*S \hat{=} (R^*S)^*$
2. $(RS+R)^*R \hat{=} R(SR+R)^*$
3. $(S(R+\varepsilon))^*S \hat{=} S((R+\varepsilon)S)^*$

Hausaufgabe 4.8 (Varianten nichtdeterministischer Automaten)

Wir betrachten einen neuen Automatentyp, genannt All-Pfad-NEA (AP-NEA). Ein solcher Automat $A = (Q, \Sigma, \delta, q_0, F)$ funktioniert wie ein NEA, mit der Ausnahme, dass ein Wort genau dann akzeptiert wird, wenn *jeder* mögliche Abarbeitungspfad in einem Endzustand endet (beim NEA genügte ein Pfad). Ein Wort $w \in \Sigma^*$ wird also genau dann akzeptiert, wenn $\emptyset \neq \hat{\delta}(q_0, w) \subseteq F$ ist.

Beweisen Sie, dass AP-NEAs genau die regulären Sprachen akzeptieren. Zeigen Sie dazu, dass für jeden NEA ein äquivalenter AP-NEA konstruiert werden kann und umgekehrt. Beweisen Sie die Korrektheit Ihrer Konstruktion.

Hinweis:

Formulieren Sie zunächst eine mathematisch präzise Beschreibung der akzeptierten Sprache eines AP-NEA und überlegen Sie sich, wie Sie bereits bekannte Resultate und Methoden der Vorlesung auf dieses Problem übertragen können.

Anmerkung: In der Version 2 des Übungsblattes wurde eine Unsauberkeit bei der Definition von All-Pfad NEAs korrigiert. Der Automat akzeptiert, wenn jeder mögliche Abarbeitungspfad in einem Endzustand endet.
Die ursprüngliche Formalisierung dieser Aussage lautete "$\hat{\delta}(q_0, w) \subseteq F$". Dies läßt auch den Fall $\hat{\delta}(q_0, w) = \emptyset$ zu, also daß ein Wort akzeptiert wird, wenn jeder Abarbeitungspfad abbricht. Um dies auszuschließen, sollte die Formulierung besser "$\emptyset \neq \hat{\delta}(q_0, w) \subseteq F$" heißen.
Der Beweis der Äquivalenz hängt von diesem Unterschied übrigens nicht ab. Man könnte genauso gut die umgangssprachliche Formulierung verwenden ohne daß sich nennenswert etwas ändert.
Genau besehen besagt die umgangssprachliche Aussage "jeder Pfad endet in einem Endzustand" sogar, daß kein Pfad an einer Stelle abbrechen darf, wo δ nicht definiert ist. Dies wird auch von der obigen Formalisierung nicht abgedeckt und ist auch nur recht umständlich zu formalisieren.