

Theoretische Informatik I

Prof. Dr. Christoph Kreitz / Thomas Rath

Universität Potsdam, Theoretische Informatik, WS 2009/10

Blatt 4 (Version 4) — Abgabetermin: **24.11.2009, 10:00 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 (Konstruktion eines NEA)

Entwerfen Sie einen ε -NEA, der die Sprache $\{w \in \{a, b\}^* \mid w = a^k \text{ mit } k = 2n \text{ oder } k = 3n \text{ für ein } n \in \mathbb{N}\}$ erkennt. Konstruieren Sie aus dem ε -NEA einen äquivalenten DEA. Verwenden Sie die optimierte (iterative) Teilmengenkonstruktion. Wieviele Zustände hätte Ihr DEA, wenn Sie alle Teilmengen der Zustände des NEA gebildet hätten?

Aufgabe 4.2 (Elemente der Sprachen regulärer Ausdrücke)

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

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

Aufgabe 4.3 (Analyse regulärer Ausdrücke)

Beschreiben Sie die Sprache der folgenden regulären Ausdrücke R und S und zeigen Sie, dass diese tatsächlich die Sprachen von R und S sind.

1. $R = (1 + 0)^*10^* + 1$
2. $S = (1 + \varepsilon)01^* + 0(1 + \varepsilon)1^*0$

Aufgabe 4.4 (Konstruktion regulärer Ausdrücke)

Geben Sie reguläre Ausdrücke für die folgenden Sprachen an.

1. $\{w \in \{0, 1\}^* \mid w \text{ enthält mindestens drei Einsen}\}$
2. $\{w \in \{0, 1\}^* \mid \text{die Anzahl der Nullen in } w \text{ ist durch 3 teilbar}\}$
3. $\{w \in \{0, 1\}^* \mid w \text{ beginnt mit 0 und } |w| \text{ ist gerade oder } w \text{ beginnt mit 1 und } |w| \text{ ist ungerade}\}$
4. $\{w \in \{0, 1\}^* \mid w \text{ enthält nicht 11}\}$

Zeigen Sie für mindestens eine dieser Sprachen, dass der von Ihnen gefundene reguläre Ausdruck tatsächlich diese Sprache repräsentiert.

Hausaufgabe 4.5 (Reguläre Ausdrücke in praktische Anwendungen)

[3 Punkte]

Reguläre Ausdrücke können in einer Anwendung dazu dienen, das gewünschte Format der Eingabe des Benutzers zu beschreiben.

Entwerfen Sie einen regulären Ausdruck, der eine Angabe in einem Literaturverzeichnis beschreibt. Hier sind zwei Beispiele für eine gültige Literaturangabe:

H. Jürgensen, L. Staiger und H. Yamasaki. Finite automata encoding geometric figures. Theoretical Computer Science, 381(1-3), S. 33-43. Elsevier, 2007.

M. Fitting und R. L. Mendelsohn. First-Order Modal Logic. Kluwer, 1998.

Gehen Sie dabei schrittweise vor. Konstruieren Sie reguläre Ausdrücke für folgende Elemente:

- Name eines Autors (regulärer Ausdruck A): besteht aus einem oder mehreren Vornamen und einem Nachnamen. Die Vornamen können mit einem Punkt abgekürzt sein.
- Liste von Autoren (AL): Name eines Autors oder mehrerer Autoren, mit Komma getrennt; vor dem letzten Autor kann ein "und" statt ein Komma stehen
- Titel (T): ein oder mehrere Wörter, die aus Buchstaben bestehen. Zwischen den Wörtern ist ein Leerzeichen.
- Titel der Zeitschrift (Z): wie Titel
- Ausgabe der Zeitschrift (N): eine Zahl, optional kann in Klammern folgen: Zahl oder Zahl-Zahl
- Seitenangaben (S): s. Zahl-Zahl oder s. Zahl
- Verlag (V): wie Titel
- Erscheinungsjahr (J): vierstellige Zahl
- Literaturangabe (R): Nach der Autorenliste folgt der Titel. Dann optional Titel und Ausgabe der Zeitschrift sowie Seitenangabe. Dann in jedem Fall: Verlag und Erscheinungsjahr. Berücksichtigen Sie die Zeichensetzung (Punkt, Komma) zwischen diesen Elementen (siehe obige Beispiele). Benutzen Sie die Variablen der bisher konstruierten regulären Ausdrücke.

Der Einfachheit halber sollen nur Kleinbuchstaben (keine Großbuchstaben) und nur die Sonderzeichen Punkt, Komma, Klammern, Bindestrich für eine Literaturangabe berücksichtigt werden. Benutzen Sie die abkürzenden Schreibweisen: $[0-9]$ für den regulären Ausdruck $(0+1+2+3+4+5+6+7+8+9)$, analog $[a-z]$ für Kleinbuchstaben. $E?$ kann für $(E + \epsilon)$ stehen, wobei E ein regulärer Ausdruck ist. Für die Leerzeichen zwischen den Wörtern (bei Titeln und Verlag) benutzen Sie den regulären Ausdruck $[: \text{ whitespace } :]$. Sonstige Leerzeichen müssen Sie nicht betrachten.

Hausaufgabe 4.6 (Konstruktion eines ϵ -NEA)

[3 Punkte]

Sei L_{bc} eine Sprache, die alle Wörter aus $\{a, b, c\}^*$ enthält, die unter den letzten vier Zeichen mindestens ein b oder ein c enthalten.

Konstruieren Sie einen ϵ -NEA A , der die Sprache L_{bc} akzeptiert. Ihr Automat A darf maximal fünf Zustände haben. Von jedem Zustand aus und auf jeden Zustand darf es maximal zwei Übergänge für ein und dasselbe Eingabesymbol geben.

Geben Sie, ausgehend von Ihrem Automaten A , einen regulären Ausdruck an, der die Sprache beschreibt, die A akzeptiert.