Prof. Dr. Franz Schweiggert — Abteilung Angewandte Informationsverarbeitung — 23.11.2005
Norbert Heidenbluth — Blatt 6

Allgemeine Informatik I für Mathematiker/Wirtschaftsmathematiker (WS 2005/2006)

Abgabetermin: 30. November 2005

Unseren Ausflug in die theoretische Informatik beenden wir auf diesem Übungsblatt mit je einer Aufgabe zu Automaten und zu regulären Ausdrücken. Letztere stellen ein unverzichtbares Werkzeug beim Suchen und Editieren großer Texte dar. Reguläre Ausdrücke können in vielen UNIX-Tools verwendet werden (z.B. im `grep` *oder im* `sed`*), und deshalb sollte Aufgabe 14 nicht als reine VIM-Aufgabe angesehen, sondern die dortige Verwendung des VIM als (mittlerweile vertrautes) "Mittel zum Zweck" verstanden werden.*

Aufgabe 12: Der "Ulmer AI1-Automat" (3 Punkte)

Zigarettenautomaten, Fahrkartenautomaten und Parkscheinautomaten kennt jeder, aber der "Ulmer AI1-Automat" ist etwas ganz Besonderes. Er hat nämlich im Wesentlichen überhaupt keine sinnvolle Funktion, kann Ihnen aber bis zu 4 Punkte auf dieses Übungblatt bringen...!

Und so sieht er aus, der "Ulmer AI1-Automat":

- $Z = \{a, b, c, d\}$
- $V = \{\text{"0"}, \text{"1"}\}$
- $\delta : Z \times V \to Z$ sei durch folgende Tabelle definiert:

δ	a	b	c	d
"0"	b	a	b	d
"1"	c	d	d	d

- $s = a$
- $E = \{a\}$

Hübsch, gell? Dann bearbeiten Sie doch jetzt bitte die folgenden Teilaufgaben zu diesem Automaten:

- Handelt es sich um einen endlichen, deterministischen Automaten?
- Stellen Sie diesen Automaten graphisch dar.
- Geben Sie eine Grammatik für die durch den Automaten definierte reguläre Sprache an.
- Geben Sie vier verschiedene Sätze an, die durch den Automaten erkannt werden.

Aufgabe 13:

Damit Ihnen diese Aufgabe kein Unglück bringt, lassen wir sie einfach weg...! Falls Sie dennoch unbedingt eine Aufgabe 13 bearbeiten möchten, dann entwerfen Sie doch mal eine Grammatik für die *Blue Ball Machine*[1]. Erwarten Sie aber bitte weder Punkte noch Hilfestellung oder gar eine Musterlösung hierfür :-))

Aufgabe 14: Reguläre Ausdrücke im VIM (5 Punkte)

Was haben Mike Krüger und Albert Einstein gemeinsam? Nun, auf was auch immer Sie jetzt tippen, auf jeden Fall haben die beiden gemeinsam, daß sie in Ulm geboren wurden.

Das macht Sie neugierig, wer noch so alles aus Ulm kommt, und deshalb haben wir auf unserem `FTP-Server` die Datei `UlmerPersonen.txt` zum Download bereitsgestellt. Sie enthält eine Liste von berühmten Persönlichkeiten, die in Ulm geboren wurden (Quelle: Wikipedia).

Zufälligerweise eignet sich diese Datei auch sogleich hervorragend, für eine Übungsaufgabe zu regulären Ausdrücken. Deshalb laden Sie sich diese Datei herunter und bearbeiten mit ihr und dem VIM[2] die folgenden Aufgaben:

1. Zunächst fällt Ihnen auf, daß es mehrere leere Zeilen in der Datei gibt. Im Zusammenhang mit der Überschrift sind diese ja noch berechtigt, mitten im Text jedoch nicht. Löschen Sie daher alle unerwünschten leeren Zeilen mit Hilfe eines regulären Ausdrucks.
2. Zu Beginn jeder Zeile steht das sogenannte "Lattenzaunsymbol". Löschen Sie dieses! Aber achten Sie darauf, daß dieses Symbol nur zu Beginn der Zeile gelöscht wird und nicht in Zusammenhang mit dem evtl. Sterbedatum, wo dieses Zeichen ebenfalls verwendet wird.
3. Sichwortartige Texte gefallen Ihnen überhaupt nicht, denn Sie wollen (seit Sie in der Vorlesung alles über Grammatiken und Sätze hörten) einen ganzen Satz haben. Deshalb möchten Sie nun die Kommata, die am Ende der ersten Zeile eines jeden Eintrags stehen, durch das Wort "war" (plus Leerzeichen – der Optik wegen) ersetzen. Außerdem ist dummerweise auch noch jeder Eintrag auf zwei Zeilen verteilt, was Ihnen ebenfalls mißfällt.

   Macht aber nix, denn Sie sind ja jetzt Regulärer-Ausdruck-Profi, so daß es Ihnen ohne Mühe gelingt, mit nur einem flockigen VIM-Substitute-Kommando

[1] URL: `http://blueballfixed.ytmnd.com/`
[2] Hier ist wieder wichtig, daß Sie statt des VI den VIM verwenden!

- die Kommata am Ende der ersten Zeile eines jeden Eintrags durch “war” (mit Leerzeichen) zu ersetzen und zugleich dabei
- jeden Eintrag einzeilig zu machen.

4. Wie es sich für einen ordentlichen Satz gehört, benötigt er nun auch einen Punkt am Ende. Fügen Sie **alle** diese Punkte mit Hilfe (nur) **eines** regulären Ausdrucks ein.
5. Suchen Sie alle Wörter, die drei aufeinanderfolgende Vokale enthalten.
6. Suchen Sie nach allen dreibuchstabigen Wörtern, die mit einem großen Vokal beginnen und keine Umlaute enthalten.
7. Die Überschrift ist irgendwie verhunzt! Bringen Sie diese mit einem knackigen VIM-Substitute-Befehl wieder in Ordnung, d.h.
    - stellen Sie die richtige Reihenfolge wieder her
    - und ersetzen Sie zugleich die Doppelpunkte durch Leerzeichen.

    Natürlich ist auch diese Aufgabe wieder nur **mit einem** Befehl und unter der Verwendung von regulären Ausdrücken zu lösen.
8. Suchen Sie nach Personen, die in den Jahren zwischen 1920 und 1939 (jeweils einschließlich) entweder geboren oder gestorben sind.

Aufgabe 15: Geheime Botschaften? (2 Punkte)

Weil Sie nun so begeistert von regulären Ausdrücken sind, beenden wir dieses Übungsblatt mit einer weiteren Aufgabe zu diesem spannenden Thema:

Erklären Sie Ihrem Tutor, welche Muster die folgenden regulären Ausdrücke finden und ggf. wodurch sie ersetzt werden.

1. `/^^/`
2. `/^[^^]/`
3. `/\<\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3}\>/`
4. `s/[ \t]*,[ \t]*/,/g`
5. `/a{5} b{,6} c{4,8}/`

Hinweis: Um diese Aufgabe nicht gänzlich unübersichtlich zu machen, ist sie **nicht** auf den VIM optimiert, denn dann müssten noch einige weitere Backslashes verwendet werden.

Viel Erfolg!