

Algorithmen und Datenstrukturen

Sommersemester 2009

12. Übungsblatt

Aufgabe 1: Suchen in Texten

Gegeben ist das folgende Suchmuster:

abcabcab

sowie die zu durchsuchende Zeichenkette:

baaabcdabcabcabed

- Berechnen Sie das Feld $next[j]$ für alle $2 \leq j \leq M$ wobei M die Länge des Suchmusters ist. (2 Punkte)
- Führen Sie mit Hilfe der berechneten $next$ -Werte eine Suche auf der gegebenen Zeichenkette durch. Stellen Sie die Suche dabei analog zum Beispiel im Skript Teil 6.4 auf Seite 12 dar. (2 Punkte)

Aufgabe 2: Kontextfreie Grammatiken

Gegeben ist die kontextfreie Grammatik G :

$G = (\Sigma, \Pi, P, S)$ mit

- $\Pi = \{S, A, B\}$
- $\Sigma = \{a, b\}$
- $P = \{$
 $S \rightarrow Ab \quad | \quad Ba,$

$$\begin{array}{l} A \rightarrow Aa \mid c, \\ B \rightarrow Bb \mid b \\ \} \end{array}$$

- a) Welche Wörter *acb,caab,bbbba,bbbb* befinden sich in der Sprache, die *G* erzeugt? Begründen Sie ihre Antwort (z.B. durch Angabe des Ableitungsbaums). (4 Punkte)
- b) Lässt sich die Sprache, die *G* erzeugt, auch durch einen regulären Ausdruck beschreiben? Falls ja, geben Sie diesen an. (2 Punkte)

Aufgabe 3: Parsergenerator Javacc

Ein Compiler führt im Wesentlichen drei Schritte aus: die lexikalische Analyse, die syntaktische Analyse und die Erzeugung des Codes. Die ersten beiden Schritte werden unter dem Begriff *Parsing* zusammengefasst. *Parser* werden allerdings kaum noch 'von Hand' erstellt, sondern man benutzt dafür Werkzeuge (Parser-Generatoren). Bekannt sind hier die Werkzeuge *lex* und *yacc*. Im Java-Umfeld wird oft mit *javacc* gearbeitet. Eine Kurz-Dokumentation von *javacc* finden Sie auf der Übungswebseite (Datei 'javacc_einfuehrung.pdf').

Wenn Sie das Programm auf Ihrem privaten PC installieren möchten, finden Sie es auf der Web-Seite: <https://javacc.dev.java.net/>.

Verwenden Sie die JavaCC Version 4.2. Unter Windows muss dabei die .zip-Datei direkt unter 'C:\' entpackt werden. Das Programm *javacc* befindet sich dann im Pfad 'C:\javacc-4.2\bin' und wird durch '*javacc.bat*' aufgerufen. Das Programm muss von der 'DOS-Eingabeaufforderung' aus aufgerufen werden. Voraussetzung ist, dass Java bereits installiert wurde.

Unter Linux können Sie das Programm in den meisten Distributionen im Paketmanager installieren und zumeist über *javacc* ausführen. Arbeiten Sie sich in die Software *javacc* ein, indem Sie die Beispiele der Kurzdokumentation nachvollziehen. Weitere Beispiele sind auch bei dem Programmpaket unter der oben genannten Web-Adresse enthalten.

Aufgabe 4: Syntaxanalyse

Ein Computer bei einer Bank empfängt per Telefon Aufträge von Kunden. Die Spracherkennungssoftware liefert dabei Zahlenwerte in der Form:

fuenf hundert und zwei und dreissig euro und acht und siebzig cent

Natürlich kann der Computer mit dieser Angabe zunächst nicht so viel anfangen.

- a) Entwickeln Sie eine Grammatik, die 'gesprochene' Zahlenwerte bis 1000 auf syntaktische Korrektheit prüft. Was sind die Terminale und die Nonterminale? (4 Punkte)

- b) Geben Sie mithilfe Ihrer Grammatik Syntaxbäume für die Zahl 'sieben hundert und zwei und siebenzig euro' und die Zahl 'drei und achtzig euro und fünf cent' an. (2 Punkte)
- c) Entwickeln Sie mithilfe von *javacc* einen Parser, der die eingegebenen Zahlenworte automatisch in die entsprechenden Zahlenwerte übersetzt. (4 Punkte)
-

Abgabetermin: Donnerstag, 02. Juli 2009, 11 Uhr Abgabekasten IFIS + Code per Mail