Prof. Franz Schweiggert
Christian Ehrhardt Abteilung Angewandte Informationsverarbeitung 1. Dezember 2003
Blatt 7

Unix-basierte Implementierung kleiner Datenbanken (WS 03/04)

Abgabetermin 11.12.2003

Join (10 Punkte)

In diesem Blatt soll ein Kommando ähnlich wie `join` (vgl. z.B. Skript Abschnitt 4.7.2 oder die Lösung zu Blatt 5) implementiert werden. Das Programm `join.pl` wird mit zwei Dateinamen als Argumente aufgerufen, wobei der **Dateiname "-" für die Standardeingabe** stehen soll. Jede der beiden Dateien enthält Zeilen mit durch Doppelpunkt getrennten Spalten. Im Gegensatz zu `join` sind die Dateien aber nicht sortiert. In beiden Dateien ist die erste Spalte eine Art Schlüssel, der aber in keiner der beiden Dateien eindeutig sein muß. Das Programm soll dann alle Paare von Zeilen aus der ersten Datei und der zweiten Datei ausgeben, deren erste Spalte übereinstimmt. Dabei sollen zunächst alle Paare ausgegeben werden, die zur ersten Zeile in der ersten Datei passen, dann alle, die zur zweiten Zeile passen usw.

Hinweise

- Jede der beiden Dateien darf nur einmal gelesen werden. Der Dateiinhalt muß dann in perl-Datenstrukturen gespeichert werden.
- Ein Aufruf, bei dem für beide Dateinamen "-" angegeben wurde ist nicht zulässig.
- Das Skript soll auch mit größeren Eingaben umgehen können. 10000 Zeilen in jeder Datei sollten durchaus möglich sein.
- Ein Filehandle mit dem man von der Standardeingabe lesen kann bekommt man mit `my $stdin = new IO::File ('<&STDIN');` .
- Auf der Homepage gibt es ein größeres Testbeispiel zum herunterladen. Ein kleines Beispiel ist unten angegeben.

Beispiel

Datei 1

```
a:1
c:3
b:2
a:5
d:blub
```

Dabei 2

```
b:2
a:10
e:fehlt
a:11
```

Ergebnis

```
a:1:a:10
a:1:a:11
b:2:b:2
a:5:a:10
a:5:a:11
```