Prof. Dr. Franz Schweiggert Abteilung Angewandte Informationsverarbeitung 18. April 2002
Ingo Melzer Blatt 1

# Software Engineering Praxis (SS 2002)

Abgabetermin 26. April 2002 vor den Übungen

## Daten, nichts als Daten ... (10 Punkte)

Um Ihren Kontostand etwas zu verbessern, akzeptieren Sie einen HiWi-Job bei den Medizinern. Sie sollen einige Diabetesdaten in einer Datenbank sinnvoll ablegen und für ein Paper zum Thema „12,6 Jahre Datenhaltung in der Medizin“ ein paar Ergebnisse liefern. Prof. Dr. sie schaffen Das Schon gibt Ihnen folgende Anweisung:

„Also, als erstes hätten wir mal die persönlichen Daten der Patienten. Die sollten auf jeden Fall zusammen abgelegt werden. Das wären vermutlich eine Nummer, Name und Vorname, Geburtsdatum, Tag der Diabetes Manifestation und das Geschlecht. Das letzte Datum muss natürlich nicht vorliegen, da hin und wieder sich auch Leute ohne Zucker zu uns verirren. Der Rest liegt aber immer vor. Der Zugriff erfolgt meistens über die Nummer das Patienten, das sollte daher schnell gehen. Das gilt auch für Zugriffe auf die folgenden Daten.

Dann haben wir noch eine Sammlung von Messungen. Den HbA1c und die zwei Blutdruckwerte kann man zusammen ablegen, da oft alle drei bestimmt werden. Pro Tag gibt es nie mehr als eine Messung. Und dann hätten wir noch einen Laborwert, der öfters ein paar Tage später kommt. Diesen sollte man wohl extra ablegen. Einer meiner Doktoranden hat diese Daten schon die letzten zwei Stunden mit der Maus in eine eigene Datei für Sie geschoben.

Ja, und dann hätten wir noch eine Sammlung von Freitexten. Diese können einmal am Tag pro Patient eingegeben werden. Die sollte man wohl auch extra ablegen.

Ich habe gehört, dass – wie war der Name nochmal – ja, die dritte Normalform bei Datenhaltung gerade modern ist, aber das schaffen Sie ja sicher locker. Ja, und Abhängigkeiten sollte man auch erkennen können.

Vielen Dank und bis nächste Woche.

Oh, nein halt. Ich hätte es fast vergessen. Ich bräuchte ja noch ein paar Abfragen.

Als erstes sollte man wohl das Manifestationsdatum so umwandeln, dass es als leer erkannt werden kann. Das sollte ja schnell gehen. So, und dann bräuchte ich eine Liste der Patientennummern von Patienten, die Diabetes haben und bei denen mindestens ein Freitext eingetragen ist. Die Liste sollte kurz sein, also bitte keine Doppelten.

Dann hätte ich noch gerne eine Liste mit Vornamen, Name und Tag von den Patienten, die am gleichen Tag das Licht des Kreissaals erblickt haben. Sortieren Sie die Ausgabe doch einfach nach dem Datum.

So, dann gibt es noch Patienten, bei denen wirklich mal am gleichen Tag der Laborwert kam, an dem auch der HbA1c gemessen wurde. Diese hätte ich auch gerne. Ach, liefern Sie mir am besten gleich eine Liste aus Vorname, Name, Datum, HbA1c und Ualbumin-Wert. Sortieren Sie doch bitte die Ausgabe nach den Namen und falls es mehrere gibt hätte ich bei jedem gerne die neuesten Daten zuerst.

Jetzt haben wir es fast. Zum Schluss brauche ich noch eine Liste meiner treusten Kunden. Bitte liefern Sie mir doch die Namen der Patienten, bei denen ich den HbA1c schon mehr als 40 mal gemessen habe. Natürlich sollte diese Liste nach dieser Häufigkeit sortiert sein. Die bekommen dann eine Bonusmessung.

Das wars jetzt aber wirklich.

Ups, Sie sind ja noch da, auf an die Arbeit."

Noch ein paar Tipps:

- Die Daten liegen unter *diab* auf dem FTP-Server.
- Im Skript steht ab Seite 16 eine Anleitung zum Anlegen einer Datenbank.
- Passen Sie *createtables* an. Zum Einfügen der Datensätze können Sie ja diese Datei kopieren und dann ab Zeile 19 bearbeiten. Einfach die Dateien vom FTP-Server einfügen und über zwei Substitutes das „insert ..." erzeugen.
- Bei der dritten Abfrage gibt es etwa 20, bei der vierten 10 und bei der letzten wieder etwa 20 Datensätze.
- Alle Abfragen natürlich in SQL.