



1. Übungszettel

(Abgabe 25.04.2008)

Einführung in Datenbanksysteme
Datenbanken für die Bioinformatik

Heinz Schweppe, Jürgen Broß, Manuel Scholz

Modalitäten des Übungsbetriebs:

- Ausgabe der Übungszettel ist stets Freitagnachmittag.
- Die Abgabe eines Übungzettels erfolgt eine Woche nach Ausgabe bis **spätestens Freitag 15:00 Uhr**.
- Verspätet abgegebene Zettel werden abgewertet (bis Montag → **halbierte Punktzahl**, ab Dienstag → **0 Punkte**).
- Von jedem Übungszettel wird **mindestens eine Aufgabe korrigiert**, welche aber nicht vorher bekannt gegeben wird/werden. Alle Aufgaben eines Übungzettels sind zu bearbeiten.
- Es werden nur die korrigierten Aufgaben in die Wertung des Übungzettels mit einbezogen.
- Mindestens **50%** der Übungspunkte sind notwendig zur Erfüllung des Kriteriums „aktive Teilnahme“. Die Gesamtwertung der Übungszettel ergibt sich aus dem arithmetischen Mittel der einzelnen Ergebnisse.
- Es dürfen maximal 2 Übungszettel mit einer Wertung von unter 20% eingereicht werden.

Übungsaufgaben

1. Aufgabe (Dateisystem vs. Datenbanksystem)

a)

Ein wesentliches Merkmal eines Datenbanksystems ist die Abstraktion von der physischen Repräsentation der Daten, sowie der Realisierung von Operationen auf diesen Daten. Als Nutzer teilt man dem Datenbanksystem anhand einer deklarativen Sprache (SQL) mit, welche Daten geliefert werden sollen und kümmert sich nicht um den eigentlichen, physischen Zugriff.

Werden Daten hingegen in einem einfachen Dateisystem gespeichert ist ein solcher deklarativer Zugriff nicht möglich. Der Nutzer muss den Zugriff selbst (imperativ) programmieren:

- Laden Sie sich von der Übungsseite eine CSV-Datei (comma separated values) herunter. Diese enthält einen Ausschnitt aus unserer Beispieldatenbank (Terra-DB), die auch in weiteren Übung zum Einsatz kommen wird. Die Datei enthält alle Städte aus der Terra-DB samt Attributen.
- Erstellen Sie ein Java-Programm, das über diese Datei iteriert und folgende Abfrage beantwortet:
Selektiere alle Städte, die mit dem Buchstaben „S“ beginnen und deren Einwohnerzahl größer als 100.000 ist und ordne die Ergebnismenge absteigend nach Einwohnerzahl.
- Das Ergebnis der Abfrage soll in der Konsole (System.out) zeilenweise ausgegeben werden.

- Liefern Sie das Programm in einer .jar-Datei ab. Die Eingabe-CSV-Datei soll **nicht** im .jar enthalten sein. Das Programm kann und soll erwarten, dass die Eingabe-CSV-Datei im Arbeitsverzeichnis liegt.

b)

- Nennen und erläutern Sie weitere Hauptmerkmale des Datenbankansatzes und diskutieren Sie auf welche Weise sich dieser vom Einsatz traditioneller Dateisysteme unterscheidet.
- Nennen und erläutern Sie die Funktionen, die ein DBMS bereitstellen sollte.

2. Aufgabe (Abstraktion in Datenbanksystemen)

- Gegeben sei folgendes Szenario:

Jede Fluglinie besitzt ein oder mehrere Flugzeuge und beschäftigt eine gewisse Anzahl Piloten und Techniker. Die Flugzeuge werden durch Hersteller, Typ und Seriennummer identifiziert. Sie werden von verschiedenen Piloten geflogen und von Technikern gewartet. Jeder Flug hat eine eindeutige Kennung, bestehend aus dem Namen der Fluglinie und einer Nummer. Ausserdem werden Abflugszeit, Ankunftszeit, Standort und Zielort gespeichert. Kunden können Flüge für mehrere Personen buchen. Jeder Kunde hat eine Kundennummer und eine Rechnungsadresse. Bei der Buchung des Fluges muss der Kunde den Namen, die Adresse und das Geburtsdatum des Fluggastes angeben. Für jede Buchung werden Preis und Sitzplatznummer erfasst. Die Buchungen eines Kunden für einen Flug sind über eine Buchungsnummer zu erreichen.

- Betrachten Sie das oben beschriebene Szenario. Geben Sie an, auf welches Modell/welche Abstraktionsschicht (extern, intern, konzeptionell) die folgenden Änderungen Einfluß haben. Begründen Sie Ihre Antworten kurz.
 - Zu allen Kunden wird die Email-Adresse gespeichert.
 - In gedruckten Kundenlisten darf nicht mehr das Geburtsdatum der Kunden angegeben sein.
 - Es wird eine neue Fluglinie gegründet.
 - Zum schnellen Zugriff wird ein Index der Kundennamen angelegt.