



1. Übungsblatt

(Testwoche: 20. - 22. April 2010)

Einführung in Datenbanksysteme
Datenbanken für die Bioinformatik

Heinz Schweppe, Katharina Hahn

Modalitäten des Übungsbetriebs:

- Ausgabe der Übungsblatt ist stets **donnerstags Nachmittag**.
- In der darauffolgenden Woche (5 – 7 Tage nach Ausgabe) findet in den Tutorien jeweils ein kurzer Test (ca. 20 Minuten) statt, in dem jeder Teilnehmer (keine Teams!) mindestens eine Aufgabe des aktuellen Blatts lösen muss. Welche Aufgaben das sind, wird jeweils zu Anfang des Tutoriums bekannt gegeben.
- Die Lösung wird von den Tutoren korrigiert und bewertet. Nur korrigierte Aufgaben werden in die Wertung des Übungsblatts mit einbezogen.
- Mindestens **60%** der Übungspunkte sind notwendig zur Erfüllung des Kriteriums der „aktiven Teilnahme“. Die Gesamtwertung der Übungsblätter ergibt sich aus dem arithmetischen Mittel der Ergebnisse der einzelnen Blätter.

Übungsaufgaben

1. Aufgabe (Dateisystem vs. Datenbanksystem)

14 Punkte

Der für die Administration der Studierendendaten verantwortliche H. Acker beschließt, das Verwaltungssystem direkt in Java zu implementieren. Von jedem Studierenden werden die *Matrikelnummer, Name, Vorname, Fachrichtung*, sowie die *Semesterzahl* gespeichert.

Das Ihnen zur Verfügung gestellten Archiv ue01.jar enthält in der Klasse `StudentAdministrationImpl.java` bereits Methoden, um eine Studierendendatei zu öffnen und zu schließen. Zum Dateizugriff wird hierbei die Klasse `java.io.RandomAccessFile` verwendet. Erweitern Sie im Folgenden die Klasse `StudentAdministrationImpl.java` um die fehlenden Methoden der Schnittstelle `StudentAdministration.java`. Greifen Sie dabei direkt über die Datei auf die Daten zu (d.h., laden Sie **nicht alle**, sondern **nur die benötigten** Datensätze in den Speicher), indem Sie sich z.B. eine `HashMap` mit der Matrikelnummer als Schlüssel auf die entsprechende Speicherstelle der Datei verwalten. Erweitern Sie die Klassen ggf. um geeignete weitere Hilfsmethoden. Entwerfen Sie vorher ein geeignetes Dateiformat, um die Daten in variabler oder fester Länge abspeichern zu können. Mit dem Testdatensatzgenerator können Sie sich geeignete Testdatensätze erzeugen lassen.

a) Implementieren Sie folgende Methoden:

- `void insertStudent(Student student);`
// Einfügen eines neuen Satzes am Ende der Datei ein
- `Student findStudentByIndex(int i)`
throws `IllegalArgumentException`;
//Liest den Datensatz des i-ten Studierenden in der Datei.
// 0 ≤ i < aktuelle Anzahl Studierender und gibt ein Objekt der Klasse Student zurück

Welche Unterschiede ergeben sich abhängig von der Entscheidung, ob die Studierendendatensätze feste oder variable Länge haben? Nennen Sie jeweils Vor- und Nachteile.

b) Implementieren Sie folgende Methoden zum Auffinden von Studierenden:

- `Student findStudentByMatrikel(int matrikel);`
// Liefert den Studierenden mit der Matrikelnummer matrikelNr.
- `Collection<Student> findStudentenByName(String name);`
// Liefert alle Studierenden mit Namen name.

c) Implementieren Sie die folgende Methode

- `Student maxStudentSemester();`
//Liefert einen Studierenden mit maximaler Semesterzahl

2. Aufgabe (Abstraktionsebenen von Datenbanksystemen)

10 Punkte

Eine Verkehrsgesellschaft, die bestimmte Bahnstrecken bedient, verwaltet Informationen über ihr Zugpersonal, sowie ihre Strecken und Züge mit Hilfe einer Datenbank. Es wird darin u.a. Informationen über die Strecken (Start- und Zielbahnhof, sowie alle Zwischenhalte, Abfahrtszeit, usw.) gespeichert und welcher Zugführer welchen Zug auf welcher Strecke wann fährt.

Geben Sie an, auf welche Abstraktionsebenen (extern, intern, konzeptionell) sich die folgenden Änderungen auswirken. Begründen Sie Ihre Antworten kurz.

- a) Auf der Druckausgabe der Streckenpläne wird jetzt zusätzlich die Haltedauer in jedem Bahnhof in Minuten angegeben.
- b) Das Unternehmen bietet neuerdings auch einen Transferservice mit PKWs auf bestimmten Strecken an. Im Gegensatz zu Zügen wird bei einem PKW die Anzahl der Sitzplätze gespeichert.
- c) Es wird eine neue Verbindung von Wanne-Eickel nach Buxtehude angeboten.
- d) Mit Hilfe eines Index soll der Zugriff auf die Abfahrtsinformationen bestimmter Bahnhöfe beschleunigt werden.
- e) Den Schaffnern werden auf ihren Dienstplänen die Namen ihrer Kollegen, die ebenfalls auf dieser Strecke eingesetzt werden, angezeigt.

3. Aufgabe (Begriffe)

8 Punkte

Beantworten Sie folgende Fragen und erläutern Sie Ihre Antwort jeweils knapp.

- a) Was versteht man unter „Logischer Datenunabhängigkeit“?
- b) Aus welchen zwei Teilen setzt sich ein Datenmodell zusammen? Welche Funktion erfüllen diese?
- c) Was ist der Unterschied zwischen einer Datenbank und einem Datenbankschema?
- d) Was versteht man unter Integrität von Daten? Geben Sie ein Beispiel an.
- e) Geben Sie zwei nicht-funktionale Eigenschaften von Datenbanksystemen an.