

Klausur Datenbanken
Wintersemester 2008/2009
Prof. Dr. Wolfgang May
21. Januar 2009, 14-16 Uhr
Bearbeitungszeit: 90 Minuten

Vorname:

Nachname:

Matrikelnummer:

Studiengang:

Bei der Klausur sind **keine Hilfsmittel** (Skripten, Taschenrechner, etc.) erlaubt. Handies müssen ausgeschaltet sein. Papier wird gestellt. Benutzen Sie nur die **ausgeteilten**, zusammengehefteten **Blätter** für Ihre Antworten. Schreiben Sie mit blauem/schwarzem Kugelschreiber, Füller, etc.; Bleistift ist nicht erlaubt.

Zum **Bestehen** der Klausur sind **45** Punkte hinreichend.

- meine Note soll mit Matrikelnummer so bald wie möglich auf der Vorlesungs-Webseite veröffentlicht werden.
- meine Note soll nicht veröffentlicht werden; ich erfahre sie dann aus Munopag/Wopag/FlexNever oder beim zuständigen Prüfungsamt.

	Max. Punkte	Erreichte Punkte
Aufgabe 1 (ER-Modell)	20	
Aufgabe 2 (Transformation ins Relationale Modell)	20	
Aufgabe 3 (SQL und Relationale Algebra)	35	
Aufgabe 4 (Indexierung und Auswertung)	15	
Summe	90	

Note:

Name:

MatNr.:

Themenstellung

Alle Klausuraufgaben basieren auf einem gemeinsamen "Auftrag": Entwerfen Sie ein Bibliotheksinformationssystem, das Bestand und Ausleihstatus *aller* Bibliotheken der Universität beinhaltet. Es wird angenommen, dass jeder Fakultät genau eine Bibliothek zugeordnet ist.

Die folgenden Informationen sollen darin abgelegt sein:

- Welche Bücher gibt es in welchen Bibliotheken?
- Welche Person hat welche Buchexemplare ausgeliehen?

Aus den folgenden Hinweisen können Sie ableiten, welche Attribute benötigt werden. (Weitere Hinweise geben auch die zu beantwortenden Anfragen in Aufgabe 3.)

Beispiele für Benutzerausweise:

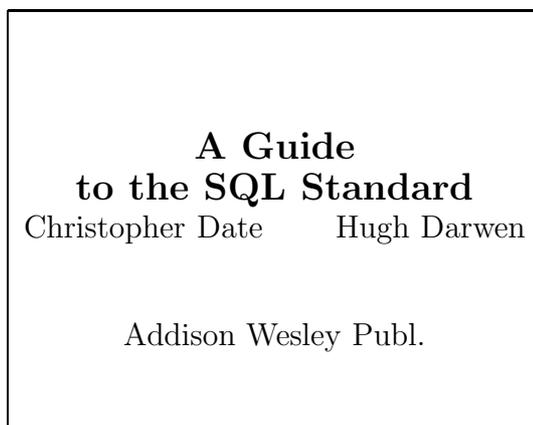
Name:	Potter
Vorname:	Harry
Benutzer-ID:	13
Statusgruppe:	Student

Name:	Einstein
Vorname:	Albert
Benutzer-ID:	5050
Statusgruppe:	Professor

In der Mathematik-Informatik-Bibliothek stehen mehrere Exemplare des folgenden Buches:

Buchdeckel:

mit Aufklebern auf dem Buchrücken:



MATHINF
DB123-1

...

MATHINF
DB123-18

(die Inventarnummern werden von jeder Bibliothek nach deren eigenem Prinzip vergeben - hier könnte es bedeuten, dass es ein Datenbankbuch ist, von dem (mindestens) 18 Exemplare vorhanden sind). Das gleiche Buch ist auch in der Geographie-Bibliothek vorhanden, wobei es dort den folgenden Aufkleber hat:

GEO
Inf42

Ausleihzettel (jede Person darf in jeder Bibliothek ausleihen):

Bibliothek: MATHINF
InvNo: DB123-17
Titel: A Guide to the SQL Standard
Ausleiher: 5050
Datum: 15-01-2009

Weiterhin soll zu jeder Person die e-mail-Adresse und an welcher Fakultät sie studiert bzw. arbeitet, abgelegt werden.

Wenn ein Buch von einem Ausleiher zurückgegeben wird, werden die Leihdaten gelöscht.

Weitere Anforderung: Es soll später jederzeit möglich sein, weitere Bibliotheken, Fächer/Fakultäten und Statusgruppen hinzuzufügen.

Aufgabe 1 (ER-Modell [20 Punkte])

Entwickeln Sie ein ER-Modell für das Szenario. Geben Sie darin die Schlüsselattribute sowie die Beziehungskardinalitäten an.

Name:

MatNr.:

Aufgabe 2 (Transformation ins Relationale Modell [20 Punkte])

- a) **Lösen Sie diesen Aufgabenteil auf dem *letzten* Blatt und trennen dieses ab (und geben es am Ende mit ab!).** Dann haben Sie dieses Blatt separat zugreifbar um später damit die Aufgaben 2b, 3 und 4 (SQL, Relationale Algebra+SQL, interne Auswertung) zu lösen.

Geben Sie an, welche Tabellen (mit Attributen, Schlüsseln etc.) Ihre Datenbank enthält (keine SQL CREATE TABLE-Statements, sondern einfach grafisch). (14 P)

Markieren Sie dabei auch Schlüssel (durch unterstreichen) und Fremdschlüssel (durch überstreichen).

Geben Sie die Tabellen mit jeweils mindestens zwei Beispieldupeln (z.B. denen, die sich aus dem Aufgabentext ergeben, und weiteren erfundenen) an.

- b) Geben Sie das CREATE TABLE-Statement für die Tabelle, in der die Ausleihdaten gespeichert sind, so vollständig wie möglich an (6 P).

Aufgabe 3 (SQL und Relationale Algebra [35 Punkte])

Verwenden Sie für diese Aufgabe die von Ihnen entworfene relationale Datenbasis. Keine der Antworten soll Duplikate enthalten.

- a) Geben Sie **eine SQL-Anfrage und einen Algebra-Ausdruck** an, die folgendes berechnen:
Welche Bücher, an denen *Hugh Darwen* mitgearbeitet hat, hat die Universität, und zu welchen Bibliotheken gehören sie? (2+2 P)
- b) Geben Sie **eine SQL-Anfrage** an, die angibt, wieviele Exemplare von "A Guide to the SQL Standard" es an dieser Uni gibt. (2 P)
- c) Geben Sie **eine SQL-Anfrage oder einen Algebra-Ausdruck** an, die/der folgendes berechnet:
Die Namen derjenigen Studenten, die ein Buch aus der Bibliothek der Fakultät an der sie studieren, ausgeliehen haben. (3 P)
- d) Geben Sie **eine SQL-Anfrage oder einen Algebra-Ausdruck** an, die folgendes berechnen:
Welche Buchtitel sind in jeder Bibliothek vorhanden? (3 P)
- e) Geben Sie **eine SQL-Anfrage oder einen Algebra-Ausdruck** an, die/der alle Buchtitel ausgibt, von denen momentan alle Exemplare verliehen sind. (5 P)
- f) Geben Sie **eine SQL-Anfrage und einen Algebra-Ausdruck** an, die die Namen aller Paare von (Student, Professor) ausgeben, die an derselben Fakultät studieren bzw. lehren, und denselben Buchtitel ausgeliehen haben. (4+4 P).
- g) Geben Sie **eine SQL-Anfrage** an, die den Titel desjenigen Buches angibt, von dem momentan am meisten Exemplare verliehen sind (bzw. mehrere Titel, falls es mehrere solche Bücher gibt)? (5 P)
- h) Ein bisschen Theorie: Gegeben sei eine Relation R mit Format (=Attributnamen) $R(A, B, C, D)$ sowie eine Relation $S(B, C)$.
 R enthalte 64.000 Tupel. S sei leer.
Welche Aussagen können Sie über die Ergebnisrelation T des Ausdrucks $R \text{ div } S$ machen (*div*: relationale Division)? Beweisen oder begründen Sie Ihr Ergebnis. (5 P)

Name:

MatNr.:

Name:

MatNr.:

Aufgabe 4 (Indexierung und Auswertung [15 Punkte])

Gegeben ist wieder das Szenario aus den Aufgaben 1-3. An alle *Studierenden*, die Bücher aus der Bibliothek der *Geographie* ausgeliehen haben, soll eine e-Mail verschickt werden. Geben Sie eine SQL- oder Algebra-Anfrage an, die die benötigten e-Mail-Adressen ausgibt. Beschreiben Sie eine sinnvolle Möglichkeit, wie diese Anfrage intern ausgewertet werden kann, und welche Indexe die Datenbank dafür enthalten sollte.

Gehen Sie bei Ihren Größenabschätzungen von folgenden Daten aus:

- 30.000 Benutzer, davon 24.000 Studenten,
- derzeit 12 Bibliotheken,
- 12.000.000 Buchexemplare,
- davon sind 360.000 Buchexemplare ausgeliehen,
- Eine Speicherseite umfasst 4096 Bytes,
- Nehmen Sie an, dass im Hauptspeichercache Indexseiten behalten werden, während normale Tabellen-Speicherungs-Seiten wenn man sie nicht iterativ durchgeht, gleich wieder ausgelagert werden.

Machen Sie bzgl. Tupelgröße plausible Annahmen basierend auf Ihren jeweiligen Tabellenschemata. Begründen Sie *kurz* und nachvollziehbar, wie Sie auf Ihre Zahlen kommen. Nehmen Sie bezüglich Werteverteilung Gleichverteilung an. Geben Sie die Anzahl der Seitenzugriffe (grobe Überschlagsrechnungen) an.

Name:

MatNr.:

Name:

MatNr.:

[Trennen Sie dieses Blatt am besten vor Beginn der Bearbeitung ab]

Lösen Sie hier Aufgabe 2a