

**Klausur Datenbanken**  
**Wintersemester 2009/2010**  
**Prof. Dr. Wolfgang May**  
**2. Februar 2010, 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 FlexNever oder beim zuständigen Prüfungsamt.

	Max. Punkte	Erreichte Punkte
Aufgabe 1 (ER-Modell)	20	
Aufgabe 2 (Transformation in das Relationale Modell)	20	
Aufgabe 3 (SQL und Relationale Algebra)	38	
Aufgabe 4 (Indexierung und Auswertung )	12	
Summe	90	

**Note:**

---

## Themenstellung: Internet-Auktionsplattform

Alle Klausuraufgaben basieren auf einem gemeinsamen "Auftrag": In der Klausur soll eine Datenbank einer Internet-Auktionsplattform, wie z.B. *ebay*, entworfen werden:

- Über alle Personen, die als Verkäufer oder als Käufer auftreten, werden einige Daten abgelegt: Name (Vor+Nachname; z.B. "Max Müller"; es kann mehrere Personen mit demselben Namen geben), eindeutige Benutzer-ID (z.B. *mmueller123*), Geburtsdatum, Ort, Strasse+Hausnummer, Land.
- Um das Angebot etwas zu strukturieren, wird jeder zu verkaufende Gegenstand einer Kategorie zugeordnet: *Computer, Bücher, CDs, Weine, Souvenirs*, etc.
- Aus rechtlichen Gründen muss jeder Benutzer für jede Kategorie, in der er als Verkäufer auftritt, angeben, ob er kommerziell oder privat verkauft. *Max Müller verkauft kommerziell Computer, sowie privat Geschirr und Bücher.*

Die zu verkaufenden Gegenstände und die Auktionen sind folgendermaßen beschrieben:

- Wenn ein Benutzer einen Gegenstand verkaufen möchte, erhält er eine ID. Dafür wird eine neue Auktion eröffnet. Der Benutzer muss den Gegenstand einer Kategorie, z.B. "Bücher" zuordnen, und eine kurze Beschreibung, z.B. "Sizilianisch Kochen", angeben.

Zusätzlich zu der Bezeichnung des Objektes kann eine textuelle Beschreibung, z.B. "guter Zustand, einzelne Fettflecken" angegeben werden. Das Ende des Auktionszeitraumes wird automatisch auf 14 Tage nach der Eröffnung gesetzt.

*Max Müller hat seit am 26.1.2010 das o.g. Buch zum Verkauf angeboten.*

- Die Auktion verläuft dann wie üblich: Bieter geben Gebote ab, die jeweils höher als das aktuelle Höchstgebot sein müssen. Wenn der Zeitraum einer Auktion beendet ist, erwirbt der letzte Bieter (der also das höchste Gebot abgegeben hat), den Gegenstand. Nach Abschluss des Geschäfts können der Verkäufer und der Käufer jeweils Bewertungen (-3,...,+3) bezüglich der Fairness des Geschäftspartners abgeben.
- Zu den laufenden Auktionen wird jeweils nur das letzte Gebot mit den notwendigen Daten gespeichert. *Das bisher letzte Gebot für das o.g. Buch wurde am 1.2.2010 von Emma Schmidt abgegeben und es beträgt 8.50E.*
- Für alle beendeten Auktionen ist abgelegt, wer was von wem wann zu welchem Preis gekauft hat. Ausserdem werden die Bewertungen gespeichert.

*Max Müller hat in einer am 15.1.2010 abgelaufenen Auktion eine Flasche mit dem Bezeichnungstext "Flasche Château de Migraine, 1954" von Bernd Brot für 256E ersteigert. Er hat ihm dafür eine Bewertung von "-3" gegeben, da Bernd Brot nicht angegeben hatte, dass es sich um eine leere Flasche handelt. Umgekehrt hat Bernd Brot Max Müller mit "-3" bewertet, da er die Flasche nicht in der Kategorie "Wein", sondern unter "Souvenirs" eingeordnet hatte.*

- Für Anfragen muss es möglich sein, zwischen abgeschlossenen und noch laufenden Auktionen zu unterscheiden.

**Aufgabe 1 (ER-Modell [20 Punkte])**

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

**Aufgabe 2 (Transformation in das 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). (12 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) Beschreiben Sie kurz, was entsprechend Ihrer Modellierung gemacht werden muss, wenn der Zeitraum einer Auktion abläuft. (2 P)
- c) Geben Sie das CREATE TABLE-Statement für diejenige Tabelle (bzw. die Tabellen), in der bei Ihnen die Daten über die in der Aufgabenstellung beschriebene abgeschlossene Auktion der Flasche “Château de Migraine” abgespeichert sind, so vollständig wie möglich an (6 P).



---

**Aufgabe 3 (SQL und Relationale Algebra [38 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 die Beschreibungen aller Bücher ausgibt, die momentan ersteigert werden können (2+2 P)
- b) Geben Sie **eine SQL-Anfrage und einen Algebra-Ausdruck** an, die die Namen aller Personen ergeben, die privat Geschirr verkaufen und nicht in Deutschland wohnen. (3+3 P)
- c) Geben Sie **eine SQL-Anfrage und einen Algebra-Ausdruck** an, die alle Paare von Namen von Personen  $(A, B)$  berechnen, so dass  $B$  und  $A$  sich bereits in beide Richtungen etwas verkauft haben, und sich dabei auch jeweils gegenseitig positiv bewertet haben: (4+4 P)
- d) Geben Sie eine **SQL-Anfrage** an, die die Namen aller Personen ausgibt, welche insgesamt für mehr als 100.000E Gegenstände kommerziell verkauft haben (4 P).
- e) Geben Sie **eine SQL-Anfrage und einen Algebra-Ausdruck** an, der die Bezeichnung und Beschreibung des Gegenstandes (oder der Gegenstände, falls es mehrere solche gibt) ausgibt, die von allen bereits abgeschlossenen Auktionen den höchsten Preis erzielt haben (3+5 P).  
(Sie dürfen **ROWNUM nicht** verwenden.)
- f) Ein bisschen Theorie: Gegeben seien zwei Relationen  $R[\bar{X}]$  und  $S[\bar{X}]$  mit demselben Format.  $R$  enthalte  $m$  Tupel,  $S$  enthalte  $n$  Tupel.
  - Betrachten Sie  $R \cup S$ . Wie groß kann die Anzahl  $k$  der Tupel in  $R \cup S$  minimal bzw. maximal (bzgl.  $m$  und  $n$ ) sein (mit Begründung) (2+2 P)?
  - Betrachten Sie  $R \bowtie S$ . Welche (genaue!) Aussage können Sie über die Anzahl der Tupel in  $R \bowtie S$  machen, wenn Sie  $k = |R \cup S|$  kennen (mit Begründung) (4 P)?



Name:

MatNr.:

---





**Aufgabe 4 (Indexierung und Auswertung [12 Punkte])**

- a) Geben Sie eine SQL-Anfrage an, die die Benutzernamen aller Personen ergibt, die in *Hamburg* wohnen und einen Gegenstand für mehr als 100.000 Euro gekauft haben (d.h., es werden nur abgeschlossene Auktionen betrachtet!). (3 P)
- b) Welche Art Index unterstützt eine entsprechende Suche auf der Spalte *Preis*? (1 P)
- c) Beschreiben Sie, wie die o.g. SQL-Anfrage möglichst effizient ausgewertet wird. Sie dürfen dabei weitere Maßnahmen aus dem Kapitel “Interne Organisation und Implementierung” annehmen, die Sie für vernünftig halten.
- Beschreiben Sie grob Ihr gewähltes Vorgehen (2 P).
  - Geben Sie eine Skizze an, wie Ihre Tabellen mit den Indexen etc., die Sie bei Ihrem Vorgehen nutzen, aussehen (2 P).
  - Geben Sie die Anzahl der Seitenzugriffe (grobe Überschlagsrechnungen) an. Rechnen Sie nur da, wo es wirklich notwendig ist (4 P).

Gehen Sie bei Ihren Größenabschätzungen von folgenden Daten aus:

- Eine Speicherseite umfasst 4096 Bytes,
- Es sind 1.000.000 Kunden, 30.000.000 abgeschlossene Auktionen und 20.000 noch laufende Auktionen gespeichert.
- 10.000 Kunden wohnen in Hamburg, und 600 Auktionen hatten einen Schlusspreis von mehr als 100.000E.
- Nehmen Sie an, dass im Hauptspeichercache Indexseiten behalten werden, während normale Tabellen-Speicherungs-Seiten, wenn man sie nicht tupelweise iterativ durchgeht, gleich wieder ausgelagert werden.



Name:

MatNr.:

---



Name:

MatNr.:

---

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

**Lösen Sie hier Aufgabe 2a**