

**Klausur Datenbanken**  
**Wintersemester 2004/2005**  
**Prof. Dr. Wolfgang May**  
**10. Februar 2004, 11-13 Uhr**  
**Bearbeitungszeit: 90 Minuten**

Vorname:

Nachname:

Matrikelnummer:

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.

	Max. Punkte	Schätzung für "4"
Aufgabe 1 (ER-Modell, Relationales Modell)	24	17
Aufgabe 2 (SQL)	20	12
Aufgabe 3 (Relationale Algebra)	21	9
Aufgabe 4 (Transaktionen)	25	10
Summe	90	48

**Note:**

Die Klausur basiert auf dem folgenden "Bank"-Szenario:

- Kunden haben einen Namen und wohnen in einer Stadt (es gibt keine zwei Kunden gleichen Namens)
- Jeder Kunde unterhält *bis zu drei* Konten, jedes Konto hat eine Kontonummer.
- Auf jedem Konto ist ein bestimmter Geldbetrag,
- Manche Kunden besitzen auf ihren Konten außerdem Aktien von Firmen.
- Die Depotdaten enthalten zu jedem Aktienpaket den Namen der Firma, die Anzahl von Aktien, das Kaufdatum und den damaligen Kaufpreis pro Aktie. Ein Kunde kann Aktien einer Firma zu unterschiedlichen Zeitpunkten und unterschiedlichen Preisen gekauft haben - hier wird angenommen, dass ein Kunde Aktien nicht zweimal am selben Tag Aktien derselben Firma kauft.
- Für jede Firma ist ihr Tätigkeitsbereich und der aktuelle Aktienkurs gespeichert.

Die Datenbank hat das folgende Schema und enthält z.B. die folgenden Werte (die Tupel-Nummern werden in Aufgabe 4 benötigt):

Kunde		Konto		
Name	Stadt	Name	KNr	KStand
Meier	Kassel	$k_1$ Meier	93	7000
Müller	Kassel	$k_2$ Müller	72	15000
Schmidt	Hamburg	$k_3$ Schmidt	18	4.80
Kunze	Hannover	$k_4$ Kunze	42	-10000
		$k_5$ Kunze	13	3.50

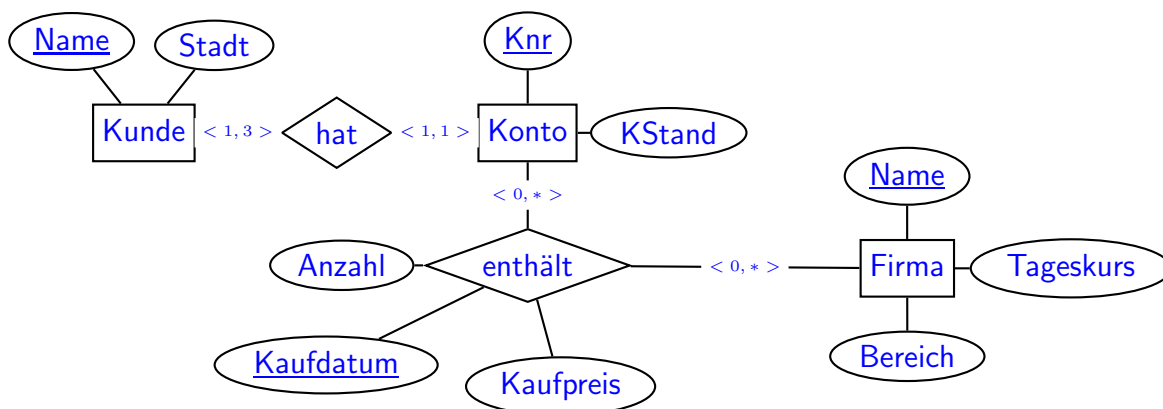
Firma		
Name	Bereich	Tageskurs
VW	Auto	26.00
Daimler	Auto	35.00
GM	Auto	1.05
SAP	IT	130.00
Nestle	Ernährung	3500
DeutscheBk	Bank	60

enthält					
	Nummer	Firma	Kaufdatum	Kaufpreis	Anzahl
$e_1$	93	VW	23.2.2000	22.00	200
$e_2$	93	Daimler	25.2.2000	83.50	60
$e_3$	93	Daimler	1.12.2001	20.00	200
$e_4$	72	GM	4.5.1998	1.20	4000
$e_5$	72	SAP	20.6.2000	640.00	10
$e_6$	18	Nestle	20.2.1960	200	10
$e_7$	18	VW	1.8.1980	18.00	300
$e_8$	42	VW	15.6.1970	1.20	2500
$e_9$	13	VW	1.7.1990	30.00	150
$e_{10}$	13	DeutscheBk	3.10.1990	40.00	100
$e_{11}$	13	Nestle	20.3.1995	2800	2

## Aufgabe 1 (ER-Modell, Relationales Modell [24 Punkte])

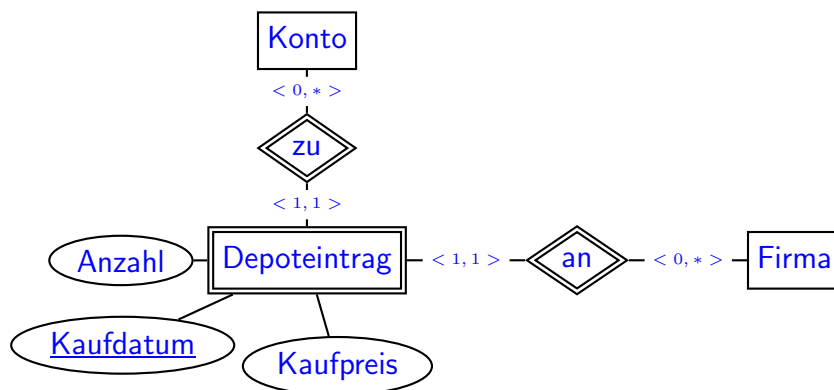
- Erstellen Sie ein ER-Diagramm für den obigen Sachverhalt (einschl. Schlüssel und Kardinalitäten) (15 P).
- Unterstreichen Sie die Schlüssel der jeweiligen Relationen (3 P):  
 Kunde(Name, Stadt)  
 Konto(Name, KNr, KStand)  
 Firma(Name, Bereich, Kurs)  
 enthält(Nummer, Firma, Kaufdatum, Kaufpreis, Anzahl)
- Überstreichen Sie in obiger Aufstellung die Fremdschlüsselattribute (2 P).
- Geben Sie zu der Tabelle “enthält” die Tabellendefinition (in SQL: CREATE TABLE ...) sowie die referentiellen Integritätsbedingungen (in SQL: FOREIGN KEY ... REFERENCES ...) an (2+2 P).  
 (die Syntax muss nicht unbedingt SQL sein, aber es muss klar sein, was Sie meinen).

### Lösung



ebenfalls richtig:

- $\langle 0,3 \rangle$  für Kunde in hat\_Konto und  $\langle 1,* \rangle$  für Konto in hat\_Konto (Gemeinschaftskonten)
- Umständlichere und fehleranfällige Modellierung: “enthält” als weak entity (z.B. “Depoteintrag”):



Schlüssel:

Kunde(Name, Stadt)

Konto(Name, KNummer, KStand)

Firma(Name, Bereich, Kurs)

enthält(Nummer, Firma, Kaufdatum, Kaufpreis, Anzahl)

```
CREATE TABLE enthaelt
( Nummer NUMBER REFERENCES Konto(KNr),
  Firma VARCHAR2(20) REFERENCES Firma(Name),
  Kaufdatum DATE,
  Kaufpreis NUMBER,
  Anzahl NUMBER,
  PRIMARY KEY (Nummer, Firma, Kaufdatum)
)
```

Andere Notation für die Referenzen z.B.:

enthaelt.nummer → konto.nr

enthaelt.firma → firma.name

Optional weitere -in der Vorlesung nicht besprochene- Bedingungen:

```
CREATE TABLE enthaelt
( Nummer NUMBER REFERENCES Konto(KNr),
  Firma VARCHAR2(20) REFERENCES Firma(Name),
  Kaufdatum DATE NOT NULL,
  Kaufpreis NUMBER NOT NULL CONSTRAINT kp CHECK (Kaufpreis > 0),
  Anzahl NUMBER NOT NULL CONSTRAINT anz CHECK (Anzahl > 0),
  CONSTRAINT enthaeltkey PRIMARY KEY (Nummer, Firma, Kaufdatum)
)
```

## Aufgabe 2 (SQL [20 Punkte])

1. Geben Sie eine SQL-Anfrage an, die die Namen aller Firmen ergibt, die Autos bauen (2P).

### Lösung

```
select name
from firma
where bereich = 'Auto';
```

2. Geben Sie eine SQL-Anfrage an, die die Menge aller Paare (Kunde, Firma) ergibt, so dass der Kunde Aktien dieser Firma besitzt (3 P).

### Lösung

```
select distinct(name, firma)
from konto, enthaelt
where konto.knr = enthaelt.nummer
```

(fehlendes distinct: - 1/2 P)

3. Geben Sie eine SQL-Anfrage an, die für jeden Kunden angibt, wieviele Aktien er insgesamt besitzt (5 P).

### Lösung

```
select name, sum(anzahl)
from konto, enthaelt
where konto.knr = enthaelt.nummer
group by name
```

Hinweis: `group by knr` ist nicht korrekt, da ein Kunde mehrere Konten haben kann.

4. Geben Sie eine SQL-Anfrage an, die alle Paare (Kunde, Firma) angibt, so dass der Kunde Aktien der Firma hat, die er teurer gekauft hat als sie heute sind (5 P).

### Lösung

```
select konto.name, firma.name
from konto, enthaelt, firma
where konto.knr = enthaelt.nummer
and enthaelt.firma = firma.name
and enthaelt.kaufpreis > firma.tageskurs
```

5. Geben Sie in natürlicher Sprache an, was die folgende Anfrage ergibt (3 P). Geben Sie weiterhin eine Zeile des Ergebnisses für die oben angegebene Beispieldatenbasis an (2 P).

```
select konto.name, sum(firma.tageskurs * enthaelt.anzahl)
from konto, enthaelt, firma
where konto.knr = enthaelt.nummer
      and enthaelt.firma = firma.name
group by konto.name
```

**Lösung** Aktueller Gesamtwert der Aktien jedes Kunden. Schmidt hat 10 Nestle-Aktien mit einem Wert von 35000 E sowie 300 VW-Aktien im Wert von je 26.00 E, also 7800 E, insgesamt 42800 E. Weitere Werte: Meier 14300E, Müller: 5500 E, Kunze (zu den beiden Konten 43 und 13) insgesamt 75900E.

### Aufgabe 3 (Relationale Algebra [21 Punkte])

1. Geben Sie einen Algebra-Ausdruck oder -Baum an, der die Menge der Namen aller Kunden ergibt, die mindestens ein Konto mit negativem Kontostand haben (2 P).

#### Lösung

$$\pi[\text{name}](\sigma[\text{KStand} < 0](\text{konto}))$$

2. Geben Sie einen Algebra-Ausdruck oder -Baum an, der die Menge der Namen aller Kunden ergibt, die Volkswagen-Aktien besitzen (4 P).

#### Lösung

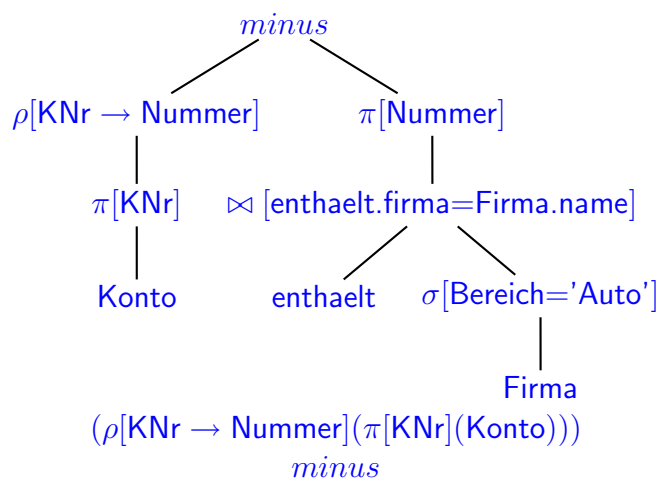
$$\pi[\text{name}](\sigma[\text{Firma} = \text{'Volkswagen'}](\text{konto} \bowtie_{[\text{KNr} = \text{nummer}]} \text{enthaelt}))$$

oder

$$\pi[\text{name}](\text{konto} \bowtie_{[\text{KNr} = \text{nummer}]} (\sigma[\text{Firma} = \text{'Volkswagen'}](\text{enthaelt})))$$

3. Geben Sie einen Algebra-Ausdruck oder -Baum an, der die Menge aller existierenden Kontonummern ergibt, die keine Automobilaktien enthalten (5 P).

#### Lösung



$$(\rho[\text{KNr} \rightarrow \text{Nummer}](\pi[\text{KNr}](\text{Konto}))) \text{ minus } (\pi[\text{Nummer}](\text{enthaelt} \bowtie_{[\text{enthaelt.firma} = \text{Firma.name}]} (\sigma[\text{Bereich} = \text{'Auto'}](\text{Firma}))))$$

Hinweis: bei *minus* und *div* muss darauf achten, dass die Signaturen geeignet aufeinander abgestimmt sind (ggf. Umbenennen).

Übung: Geben Sie einen Ausdruck für die folgende sehr ähnlichen Anfrage an (die in der Klausur oft als Antwort angegeben wurde):

- alle Konten, die (unter anderem) Aktien von Firmen enthalten, die nicht im Automobilbereich tätig sind (hier ist das *minus* an einer anderen Stelle).

4. Geben Sie einen Algebra-Ausdruck oder -Baum an, der die Namen aller Kunden angibt, die Aktien *aller in der Datenbank enthaltenen Automobilfirmen* besitzen (5 P).

#### Lösung

$$\pi[\text{Name}, \text{Firma}](\text{konto} \bowtie_{[\text{KNr} = \text{nummer}]} \text{enthaelt}) \text{ div } \rho[\text{Name} \rightarrow \text{Firma}](\pi[\text{Name}](\sigma[\text{Bereich} = \text{'Auto'}](\text{Firma})))$$

(auch hier: vor der Division passend umbenennen.)

Hinweis: dazu müssen ggf. alle Konten eines Kunden zusammen betrachtet werden.

Übung: Geben Sie einen Algebra-Ausdruck oder -Baum an, der die Namen aller Kunden angibt, die (mindestens) ein Konto haben, auf dem Aktien *aller in der Datenbank enthaltenen Automobilfirmen* liegen (hier ist das *div* an einer anderen Stelle).

5. Gegeben seien zwei nichtleere Relationen  $R(A, B)$  und  $S(B, C)$ . Ist für diese Relationen die Gleichung  $\pi[A, C](R \bowtie S) = (\pi[A](R)) \bowtie (\pi[C](S))$

- gültig,
- abhängig von den Inhalten der Relationen mal gültig, mal nicht gültig,
- immer ungültig.

Begründen Sie Ihre Entscheidung (5 P).

**Lösung** Die Gleichung ist abhängig von den Inhalten der Relationen mal gültig, mal nicht gültig.

Bei der linken Seite wird über  $B$  gejoint, während auf der rechten Seite ein kartesisches Produkt gebildet wird. Damit ist die Gleichung genau dann gültig, wenn es für jedes  $A$ - $C$ -Paar  $(a, c)$  mindestens einen  $B$ -Wert  $b$  gibt, so dass  $(a, b) \in R$  und  $(b, c) \in S$  sind. (Da die relationale Algebra mengenorientiert ist, spielen eventuelle Duplikate keine Rolle).



#### Aufgabe 4 (Transaktionen [25 Punkte])

Sonntag abend, 13.1.2005: Schmidt muss eine Rechnung in Höhe von 10000 E bei Müller bezahlen. Dazu gibt er eine online-Überweisung für Montagmorgen in Auftrag. Da er schon kein Geld mehr auf dem Konto hat, entschließt er sich, Aktien zu verkaufen und vergibt einen online-Auftrag, am nächsten Morgen 3 Nestle-Aktien limitiert mit 3400 E zu verkaufen (da dies unter dem Tageskurs liegt, ist er sich ziemlich sicher, dass der Verkauf klappt). Müller, wissend dass er am nächsten Tag 10000 E von Schmidt überwiesen bekommt, und auch noch Geld auf dem Konto hat, will diese neu anlegen, und schickt einen Kaufauftrag für 3 Nestle-Aktien, limitiert auf 3400 E ab.

Wie der Zufall will, liegen am Montagmorgen keine weiteren Kauf- oder Verkaufsaufträge für Nestle vor, und es geschehen die folgenden Dinge (jeweils mit Angabe der entsprechenden Tupel in den Tabellen):

$T_1$  Überweisung von 10000 E vom Konto von Schmidt (Tupel  $k_3$ ) auf das Konto von Müller ( $k_2$ ).

$T_2$  Ausbuchung von 3 Nestle Aktien aus dem Depot von Schmidt ( $e_6$ ) und Gutschrift von 10200 E auf dem Konto von Schmidt ( $k_3$ ).

$T_3$  Einbuchung von 3 Nestle Aktien in das Depot von Müller ( $e_{12}$ , neues Tupel) und Abbuchung von 10200 E von dem Konto von Müller ( $k_2$ ).

... und jetzt zur Aufgabe:

1. Komplettieren Sie die Transaktionen  $T_1$  und  $T_2$ , die die jeweiligen Tupel  $k_2$ ,  $k_3$ , und  $e_6$  der Datenbank lesen/schreiben; die dritte Transaktion (die u.a. das Tupel  $e_{12}$  schreibt), ist angegeben (4 P):

- $T_1 = R_1 k_3; k_3 = k_3 - 10000; W_1 k_3; R_1 k_2; k_2 = k_2 + 10000; W_1 k_2$
- $T_2 = R_2 e_6; e_6 = e_6 - 3; W_2 e_6; R_2 k_3; k_3 = k_3 + 10200; W_2 k_3$
- $T_3 = W_3 e_{12}; R_3 k_2; k_2 = k_2 - 10200; W_3 k_2$

2. Geben Sie die Tupel  $k_2$ ,  $k_3$ ,  $e_6$  und  $e_{12}$  nach korrekter Ausführung aller Transaktionen an (2 P).

#### Lösung

$k_2 = \text{Konto}(\text{Müller}, 72, 14800)$

$k_3 = \text{Konto}(\text{Schmidt}, 18, 204.80)$

$e_6 = \text{enthaelt}(18, \text{Nestle}, 20.2.1960, 200, 7)$

$e_{12} = \text{enthaelt}(72, \text{Nestle}, 14.1.2005, 3400, 3)$

3. Geben Sie einen beliebigen Schedule  $S$  an, der die drei Transaktionen korrekt ausführt. (2 P)

#### Lösung Beliebige serielle Schedules, z.B. $T_1 T_2 T_3$ .

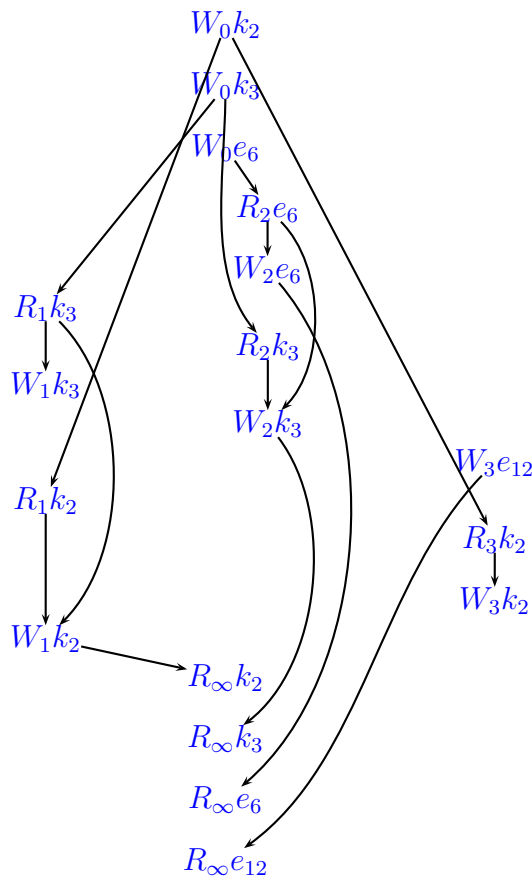
4. Geben Sie einen Schedule  $S'$  an, bei dem die drei Transaktionen nicht korrekt ausgeführt werden. Betrachten Sie dabei nur die Lese- und Schreibaktionen, geben Sie dabei jeweils die gelesenen/geschriebenen Werte an. (5 P)

**Lösung** Der folgende Schedule z.B. ignoriert die Abbuchungen von  $k_2$  und  $k_3$ :  
 $R_2e_6(10); W_2e_6(7); R_1k_3(4.80); R_2k_3(4.80); W_1k_3(-9995.20); W_2k_3(10204.80);$   
 $W_3e_{12}(\dots); R_1k_2(15000); R_3k_2(15000); W_3k_2(4800); W_1k_2(25000)$

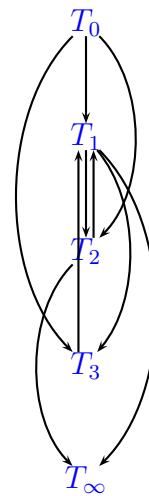
5. Geben Sie den Dependency-Graphen Ihres Schedules  $S'$  an (berücksichtigen Sie dabei auch die Transaktionen  $T_0$  und  $T_\infty$  wie in der Vorlesung) (4 P).
6. Geben Sie den Konflikt-Graphen Ihres Schedules  $S'$  an (4 P).
7. Woran kann man erkennen, dass  $S'$  nicht serialisierbar ist? (1 P)

**Lösung**

Dependency-Graph:



Conflict Graph:



Der Conflict Graph besitzt mehrere Zyklen. Dies ist allerdings nur dann ein hinreichendes Kriterium, dass der Schedule nicht serialisierbar ist, wenn es keine *Blind Writes* gibt. Im vorliegenden Fall gibt es ein Blind Write ( $W_3e_{12}$ ). Da der Zyklus die Transaktion  $T_3$ , die das Blind Write ( $e_{12}$ ) durchführt, nicht beinhaltet, ist  $S'$  nicht serialisierbar.

Hinweis: kein Eintrag  $W_0e_{12}$ ,  $W_3e_{12}$  ist ein *blind write*.

8. Zeigen Sie, dass unter Anwendung des 2-Phasen-Locking-Protokolls Ihr Schedule  $S'$  nicht zustandekommen kann. (3 P)

**Lösung** Wichtig sind hierbei die Write-Locks, weniger die R-Locks (Lesen können mehrere Transaktionen gleichzeitig; nur wenn jemand ein W-Lock hat/haben will, dürfen keine R-Locks vergeben werden/sein).

$T_1$  müsste nach dem Schreiben von  $k_3$  das W-Lock abgeben, sich also bereits in der Abbauphase befinden. Umgekehrt kann  $T_1$  aber das W-Lock auf  $k_2$  erst nach der vorletzten Aktion (in der  $T_3$   $k_2$  noch schreibt) bekommen, sich also dort noch in der Aufbauphase befinden. Dies ist nicht möglich.