Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Prof. H.-J. Schek *A. Feßler, U. Röhm, H. Schuldt, R. Weber*

Informationssysteme IS-G SS 1997

Prüfungsvorbereitung

Aufgabe 1: Relationale Algebra (Klausur WS96/97, Aufg 1)

Gegeben sei das folgende Schema einer Datenbank mit Sportdaten. (Nehmen Sie an, Sportlernamen seien eindeutig. Das Datumsformat sei 'JJMMTT', d.h. '970224' bezeichnet den 24.2.97.)

`Sportler (Name, Geburtsdatum, Land, Sponsor)`
Z.B: ('Hans Müller', 450312, 'Liechtenstein', 'Trabant')

`Weltrekord (Sportart, Datum, Name, Wert)`
Z.B.: ('Hallenhalma', 700424, 'Hans Müller', '45 Stunden')

`Sponsor (Name, Land, Jahresumsatz)`
Z.B.: ('Trabant', 'Deutschland', 0)

Schreiben Sie für die folgenden Fragen jeweils Ausdrücke der Relationenalgebra und SQL-Ausdrücke. Wenn eine Frage nicht in Relationenalgebra formulierbar ist, so begründen Sie, warum. Sie dürfen Viewdefinitionen verwenden, wenn dies die Lösung einer Teilaufgabe vereinfacht.

a) Nennen Sie die Weiten aller Weltrekorde im Weitsprung, die vor 1972 aufgestellt wurden.

b) Wer ist der Sponsor des Sportlers, der am 1.6.1972 den Weltrekord im Hochsprung erreicht hat?

c) Welche Weltrekordhalter haben mit einem ihrer Sponsoren entweder den Namen oder das Herkunftsland gemeinsam?

d) Alle Sponsoren, die Weltrekordler unter Vertrag haben oder hatten, die zum Zeitpunkt ihres Rekordes älter als 35 Jahre alt waren?

e) Welcher Sponsor hat die meisten Weltrekordler unter Vertrag?

Aufgabe 2: Autorisierung (Klausur WS95/96, Aufg 3)

Gegeben sei die Relation `R(A, B, C, D, E)`. Nehmen Sie an, dass Sie Eigentümer von R sind.

a) Wie gehen Sie in SQL vor, wenn Sie nicht die gesamte Relation, sondern nur `R'(A, B, C)` mit allen Rechten der Öffentlichkeit zugänglich machen wollen?

b) Ihr Chef verlangt, dass die Relation `R'` materialisiert werden soll. Wie gehen Sie in SQL vor?

c) Ihr Chef verlangt, dass Einfügen und Löschen uneingeschränkt in `R` und `R'` möglich sein sollen, ohne dass dabei inkonsistente Daten gelesen werden können. Welche Trigger müssen Sie einführen?

Aufgabe 3: Embedded SQL (Klausur WS94/95, Aufg 2)

Gegeben sei eine relationale Datenbank einer Bank mit dem folgenden Datenbankschema:

```
Konto     (KNr, Inhaber, Kontostand, Geschäftsstelle)
Buchungen (Monat, Tag, Zeit, KNr, Betrag, Text)
```

Einmal im Monat sollen für alle Konten einer Geschäftsstelle die Kontoführungsgebühren berechnet werden und den jeweiligen Konten belastet werden. Die monatlichen Gebühren berechnen sich folgendermassen:

Fr. 2.00 * Anzahl der Buchungen	falls Anzahl der Buchungen ≤ 5.
Fr. 5.00 + Fr. 1.00 * Anzahl der Buchungen	falls $5 <$ Anzahl der Buchungen ≤ 50.
Fr. 60.00	falls Anzahl der Buchungen > 50.

Schreiben Sie für diesen Zweck ein Embedded–SQL–Programm, welches eine Geschäftsstellennummer und einen Monat einliest und für alle Konten dieser Geschäftsstelle die entsprechenden monatlichen Gebühren von den Konten abbucht. Die Abbuchung der Gebühren ist selbst wiederum eine Buchung, die aber bei der Gebührenberechnung nicht mitzählt. Das Programm soll für jedes Konto jeweils eine Transaktion erzeugen. Gehen Sie davon aus, dass nach einem COMMIT WORK Cursor-Positionen weiterhin gültig sind.

Hinweis: Die Anweisungen des Wirtsprogramms sollen in Pseudocode–Notation skizziert werden; alle SQL–Anweisungen sollen voll formuliert werden.

Aufgabe 4: RSV-Berechnung

Gegeben sei eine Kollektion von multimedialen Dokumenten $\{d_0, \ldots, d_6\}$. Die Dokumente seien bereits in einem Indexierungsschritt — z.B. durch ein OCR-System, ein Spracherkennungssystem oder ein Textanalyseprogramm — in eine Sequenz von Indexierungsmerkmalen umgewandelt worden. In dieser Aufgabe nehmen wir vereinfachend an, dass Indexierungsmerkmale eindeutig identifiziert werden können. In der Kollektion treten 10 verschiedene Merkmale $\{\varphi_0, \ldots, \varphi_9\}$ auf.

$$
\begin{aligned}
d_0 &=< \varphi_0, \varphi_8, \varphi_0, \varphi_8, \varphi_7, \varphi_0 > \\
d_1 &=< \varphi_0, \varphi_2, \varphi_0, \varphi_8, \varphi_7 > \\
d_2 &=< \varphi_1, \varphi_0, \varphi_3, \varphi_7, \varphi_1, \varphi_0 > \\
d_3 &=< \varphi_0, \varphi_4, \varphi_7, \varphi_4, \varphi_7, \varphi_0 > \\
d_4 &=< \varphi_7, \varphi_2, \varphi_3, \varphi_4, \varphi_0 > \\
d_5 &=< \varphi_0, \varphi_7, \varphi_0, \varphi_2, \varphi_3, \varphi_2, \varphi_7, \varphi_5, \varphi_0, \varphi_9, \varphi_3, \varphi_2, \varphi_4, \varphi_7 > \\
d_6 &=< \varphi_0, \varphi_6, \varphi_7 >
\end{aligned}
$$

Eine Anfrage sei gegeben durch $q =< \varphi_2, \varphi_3, \varphi_4, \varphi_7 >$.

a) Berechnen sie von jedem Merkmal φ_i sowohl die Merkmalshäufigkeiten $ff(\varphi_i, d_j)$ in jedem Dokument d_j als auch die Dokumentenhäufigkeit $df(\varphi_i)$ und die inverse Dokumentenhäufigkeit

$$idf(\varphi_i) := \log_2 \left(\frac{n+1}{df(\varphi_i)+1} \right)$$

in der Kollektion, wobei n die Anzahl Dokumente in der Kollektion bezeichnet.

b) Berechnen Sie mit Hilfe der $ff \cdot idf$–Gewichtung sowohl die Dokumentenvektoren $\vec{d_j}$ als auch den Anfragevektor $\vec{q}$.

c) Bestimmen Sie die Rangliste der Dokumente bezüglich der gegebenen Anfrage unter Verwendung des RSV, wie in der Vorlesung besprochen.

Aufgabe 5: Relationaler Entwurf

Als Datenbankprofi sollen Sie ein altes, dateibasiertes Absatz-Informationssystem auf ein relationales Datenbanksystem umstellen. Alles, was Sie von Ihrem Abteilungsleiter bekommen, ist die alte Dateistruktur:

`ABSATZ (knr, kname, plz, pf, kums, og, geb, bra, ktyp, umsz, ver)`

Ihr Kollege B. Estseller aus der Verkaufsabteilung kennt die Bedeutung der Abkürzungen und die Zusammenhänge. Nach einem Interview haben Sie die folgenden Funktionalabhängigkeiten erhoben:

```
{  knr      ⟶ kums kname plz pf bra
   plz      ⟶ og geb
   bra geb  ⟶ ver
   og geb   ⟶ ktyp
   ver      ⟶ geb
   ktyp     ⟶ umsz }
```

Dabei bedeuten

`knr`	Kundennummer	`plz`	Postleitzahl
`pf`	Postfach	`kums`	Umsatzvolumen
`og`	Ortsgrösse	`geb`	Gebiet
`bra`	Branche	`ktyp`	Kundentyp
`umsz`	Umsatzzielvorgabe	`ver`	Verkäufer(-name)

a) Zeigen Sie anhand von jeweils einer Einfüge-, Lösch- und Änderungsanomalie, warum obiger Entwurf schlecht ist.

b) Führen Sie schrittweise eine verlustfreie Zerlegung von `ABSATZ` in BCNF durch. Geben Sie vor jedem Zerlegungsschritt an, in welcher Normalform die jeweiligen Relationenschemata sind und begründen Sie dies. Streben Sie bei der Zerlegung eine Abhängigkeitsbewahrung an. Überprüfen Sie Ihre Zerlegung in dieser Hinsicht und machen Sie eine Angabe über die Art des Vorgehens, um zu einer möglichst abhängigkeitsbewahrenden Zerlegung zu gelangen.

Aufgabe 6: Normalisierung (Klausur WS96/97, Aufg 6)

Eine Datenbank bestehe aus folgenden drei Relationen R, S und T und den zugehörigen Funktionalabhängigkeiten F_R, F_S und F_T:

`R(a, b, c, d)`	$F_R = \{b \to c, cd \to a, a \to bcd\}$
`S(c, e, f, g)`	$F_S = \{ef \to g, ce \to fg\}$
`T(f, g, h, i)`	$F_T = \{fg \to hi, g \to h\}$

a) Bestimmen Sie jeweils für jede der drei Relationen R, S und T sämtliche Schlüsselkandidaten.

b) Untersuchen Sie für jede der drei Relationen R, S und T ob sie sich in 3NF oder BCNF befinden (Begründung).

c) Zerlegen Sie die Relation R so, dass die Ergebnisrelationen in BCNF sind. Ist diese Zerlegung abhängigkeitsbewahrend ? Begründen Sie Ihre Antwort.

d) Können die drei Relationen R, S und T durch eine zweimalige Anwendung des BCNF-Zerlegungsalgorithmus aus der Universalrelation U mit

`U(a, c, e, f, g, b, d, h, i)`

und den Funktionalabhängigkeiten $F_U = (F_R \cup F_S \cup F_T)$ entstanden sein, ohne dass dabei noch weitere Relationen entstanden sind? (Begründen Sie Ihre Antwort)

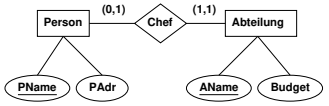
Aufgabe 7: Modellierungsvarianten

Hans Meier hat die Aufgabe, einen Entwurf für die Personalverwaltung eines grösseren Unternehmens zu erstellen. Dabei sollen u.a. die folgenden Integritätsbedingungen modelliert werden:

- Jede Abteilung hat genau einen Chef
- Eine Person kann höchstens der Chef einer Abteilung sein

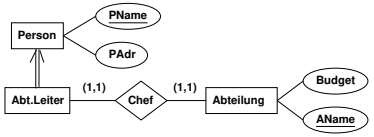
Ausgestattet mit den Erfahrungen der Vorlesung *Informationssysteme* ist Hans Meier gleich so motiviert, dass er die Situation auf drei verschiedene Arten im ERM modelliert hat:

(i)



(ii)

(iii)

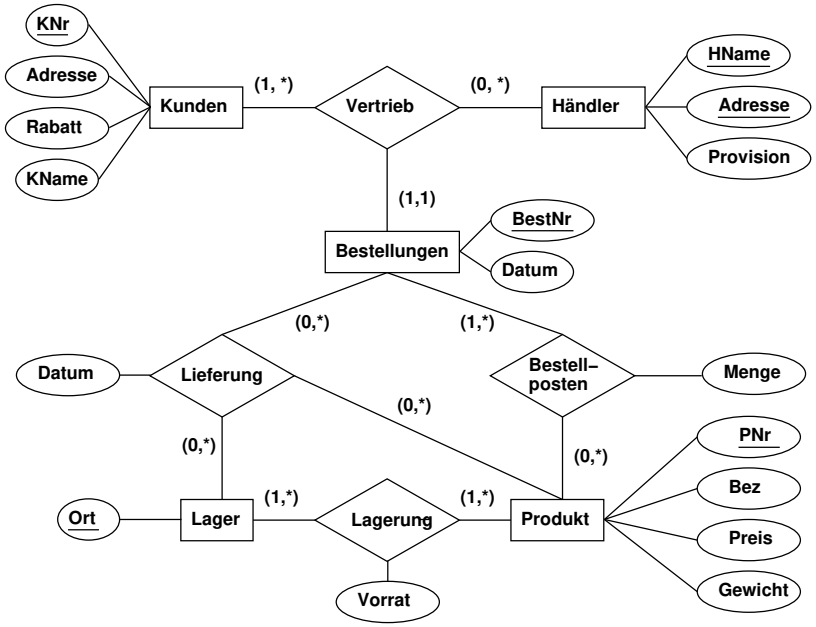


Vergleichen Sie diese drei Modellierungsvarianten:

a) Erfüllen diese die geforderten Integritätsbedingungen? Begründen Sie!

b) Welche Variante würden Sie als die Beste bezeichnen? Begründen Sie!

Aufgabe 8: Transformation ERM nach Relationenmodell

Betrachten Sie die Beziehung *Vertrieb* im nachfolgenden ERM–Diagramm. Formulieren Sie sprachlich die Bedeutung dieser Beziehung und überlegen Sie sich eine bessere Modellierungsalternative.

Transformieren Sie dann das ursprüngliche ERM–Diagramm in ein Schema des Relationenmodells. Geben Sie Schlüsselkandidaten und Fremdschlüssel an.