

Datenbanksysteme (2h)	schriftliche Einzelprüfung	24.04.2009	1
-----------------------	----------------------------	------------	---

### Aufgabe 1 [Relationale Abfragen: 30 Punkte]

Gegeben ist folgendes Relationenschema:

kunde (knr, vorname, nachname, geschlecht, alter, land)

artikel (artnr, name, beschreibung, gewicht, preis)

firma (fnr, name, strasse, plz, ort, land)

bestellung (bestellnr, *fnr*, *knr*, datum)

position (bestellnr, *artnr*, anzahl)

*bestellung.fnr* ◊ *firma*, *bestellung.knr* ◊ *kunde*, *position.bestellnr* ◊ *bestellung*, *position.artnr* ◊ *artikel*  
 artikel.gewicht in Gramm (g), artikel.preis in Euro (EUR)

Formulieren Sie die folgenden Abfragen (a, b, c) in **Relationenalgebra**:

- (3 Punkte) Ermitteln Sie Firmennummer und Name aller Firmen, die in Österreich über eine Niederlassung in Wien oder Linz verfügen.
- (4 Punkte) Ermitteln Sie Ort und Land der Firma mit der höchsten (größten) Postleitzahl.
- (5 Punkte) Ermitteln Sie die Namen aller Artikel, die bei zumindest 2 Firmen aus unterschiedlichen Ländern bestellt wurden.

Formulieren Sie die folgenden Abfragen (d, e, f, g) in **SQL99**:

- (3 Punkte) Ermitteln Sie Firmennummer und Name aller Firmen, die in Österreich über eine Niederlassung in Wien oder Linz verfügen.
- (4 Punkte) Ermitteln Sie wie viele Artikel insgesamt (alle Bestellungen zusammen genommen) bei der Firma Amazonas bestellt wurden.
- (5 Punkte) Ermitteln Sie die Namen aller Artikel, die am gleichen Tag (datum) von zumindest 5 unterschiedlichen Kunden bestellt wurden.
- (6 Punkte) Ermitteln Sie alle gemeinsamen Kunden der Firmen Winzigweich und Apfel, die zwischen (inkl.) 20 und 25 Jahre alt sind und aus Polen stammen.

### Aufgabe 2 [Query Optimierung: 30 Punkte]

Gegeben ist folgende Abfrage in relationaler Algebra, basierend auf den beiden Relationenschemata  $R_1(A, B, C, D, E)$ ,  $R_2(F, G, H, I)$ :

$$\pi_{C,E,H}(\pi_{A,C,E,F,H,I}(\sigma_{(B>=1 \wedge D=2 \wedge G<=3)}(R_1 \bowtie_{(B=G)} R_2)))$$

- (4 Punkte) Stellen Sie den Ausdruck grafisch dar.
- (12 Punkte) Führen Sie eine heuristische Optimierung mit Hilfe der in der Vorlesung angegebenen Äquivalenzumformungsregeln durch, mit dem Ziel den Abarbeitungsaufwand zu minimieren.  
 Stellen Sie den umgeformten Ausdruck grafisch dar und geben Sie die Reihenfolge der verwendeten Äquivalenzumformungsregeln an.  
 Begründen Sie Ihre Vorgangsweise.
- (14 Punkte) Nehmen Sie an, dass die Relation  $R_1$  7000 und die Relation  $R_2$  4000 Datensätze enthält, wobei die Blockgröße für beide Relationen 10 ist.

Für den Join wird das Nested-Loop Verfahren verwendet (Memorygröße 1 Block) und die Selektivität der Selektionen ist jeweils 1/10 (Annahme der Unabhängigkeit).

Geben Sie die Kosten (Anzahl von Blockzugriffen) für

- den ursprünglichen Ausdruck und
- den umgeformten Ausdruck an.

Nehmen Sie an, dass die Abarbeitung der Ausdrücke Pipelining nutzt.

### Aufgabe 3 [Formaler Datenbankentwurf: 20 Punkte]

Gegeben ist folgende funktionale Abhängigkeit:

$$RS_1 = (\{A, B, C, D, E, F\}, \{AB \rightarrow CDE, D \rightarrow ACF, AD \rightarrow E\})$$

$$RS_2 = (\{A, B, C, D, E, F, G\}, \{BD \rightarrow AEF, F \rightarrow AC, BF \rightarrow DE\})$$

Datenbanksysteme (2h)	schriftliche Einzelprüfung	24.04.2009		2
-----------------------	----------------------------	------------	--	---

- a. (5 Punkte) Geben Sie für  $RS_1$  die minimale Überdeckung der funktionalen Abhängigkeiten an.
- b. (5 Punkte) Bestimmen Sie für  $RS_2$  alle Schlüsselkandidaten.
- c. (5 Punkte) In welcher maximalen Normalform befindet sich  $RS_2$ ? Begründen Sie Ihre Aussage.

Gegeben ist folgende funktionale Abhängigkeit:

$$RS_3 = (\{A, B, C, D, E, F, G\}, \{AC \rightarrow F, A \rightarrow BE, B \rightarrow C, AD \rightarrow G, D \rightarrow AE\})$$

- d. (5 Punkte) Stellen Sie mittels verbund- und abhängigkeitstreuer Zerlegung (Synthesealgorithmus) für  $RS_3$  die dritte Normalform her.

*Anmerkung: Schlüssel der ursprünglichen Relation ist D.*

**Aufgabe 4 [Einführung und Motivation: 10 Punkte]**

Nennen und beschreiben Sie **fünf** Vorteile der Informationsverwaltung bei Verwendung eines DBMS anstatt eines Dateisystems.

**Aufgabe 5 [Relationales Modell: 10 Punkte]**

- a. (5 Punkte) Was versteht man unter der referentiellen Integrität im Zusammenhang mit Datenbanken?
- b. (5 Punkte) Wie können Abhängigkeiten die sich durch die referentielle Integrität ergeben in DDL/SQL ausgedrückt werden? Geben Sie unter Verwendung von DDL/SQL ein einfaches Beispiel an.