

Datenbanksysteme (2h)	schriftliche Einzelprüfung	26.05.2010		1
-----------------------	----------------------------	------------	--	---

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

Gegeben ist folgendes vereinfachtes Relationenschema:

konferenz (kid, name, jahr, land, stadt, kategorie)

autor (aid, vorname, nachname, gebdatum)

teilnahme (kid, aid, gebuehr)

FK: teilnahme.kid  $\diamond$  konferenz, teilnahme.aid  $\diamond$  autor

konferenz.kategorie IN {'A', 'Ae', 'B', 'Be', 'C'}

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

- (3 Punkte) Ermitteln Sie Name, Jahr und Land aller Konferenzen die zwischen 2000 und 2005 (jeweils inklusive) statt fanden und der Kategorie 'A' oder 'Ae' zugeordnet sind.
- (4 Punkte) Ermitteln Sie Vor- und Nachname des ältesten Autoren.
- (5 Punkte) Nennen Sie alle Paare von Konferenzen, die im gleichen Jahr und am selben Konferenzstandort (Land+Stadt) abgehalten wurden.

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

- (3 Punkte) Ermitteln Sie Name, Jahr und Land aller Konferenzen die zwischen 2000 und 2005 (jeweils inklusive) statt fanden und der Kategorie 'A' oder 'Ae' zugeordnet sind.
- (4 Punkte) Ermitteln Sie Vor- und Nachname des ältesten Autoren.
- (5 Punkte) Ermitteln Sie für alle Konferenzen mit dem Wort 'Cloud' im Namen den Konferenznamen und die Summe aller durch die teilnehmenden Autoren erhobenen Konferenzgebühren.
- (6 Punkte) Ermitteln Sie die Namen aller Autoren, die bereits bei einer Konferenz in 'England' als auch in 'Australien', jedoch noch nie bei einer Konferenz in 'China' teilgenommen haben.

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

Gegeben ist folgende Abfrage in relationaler Algebra, basierend auf den beiden Relationenschemata  $R_1(R, A, J, K, U, M)$  und  $R_2(B, Y)$ :

$$\pi_{R,U,B,Y}(\pi_{R,A,K,U,M,B,Y}(\sigma_{(R>=2 \wedge J<4 \wedge M=6 \wedge Y<=8)}(R_1 \bowtie_{(A=B)} R_2))))$$

- (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  $R_1$  400 und die Relation  $R_2$  300 Datensätze enthält, wobei die Blockgröße für beide Relationen 20 ist.  
Für den Join wird das Nested-Loop Verfahren verwendet (Memorygröße 1 Block) und die Selektivität der Selektionen ist jeweils  $\frac{1}{2}$  (Annahme der Unabhängigkeit).  
Geben Sie die Kosten (Anzahl von Blockzugriffen) für
  - den ursprünglichen Ausdruck und
  - den umgeformten Ausdruck an.

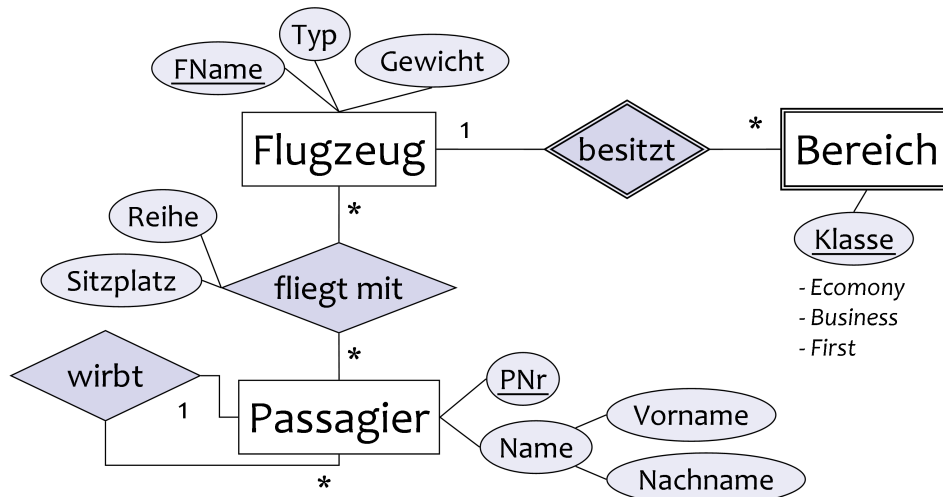
Nehmen Sie an, dass die Abarbeitung der Ausdrücke Pipelining nutzt.

### Aufgabe 3 [Formaler Datenbankentwurf: 15 Punkte]

Gegeben ist folgende funktionale Abhängigkeit:

$$RS = (\{M, E, L, B, O, U, R, N\}, \{LO \rightarrow U, BEL \rightarrow OU, O \rightarrow E, N \rightarrow L, B \rightarrow NO, M \rightarrow EL, OE \rightarrow UN\})$$

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



#### Aufgabe 4 [Relationenmodell und Datenbanksprachen: 15 Punkte]

- a. (5 Punkte) Ordnen Sie den folgenden zehn Komponenten des Entity-Relationship Diagramms den jeweils richtigen Typ zu:

'Flugzeug' ...	ist ein/eine _____
'wirbt' ...	ist ein/eine _____
'PNr' ...	ist ein/eine _____
'Name' ...	ist ein/eine _____
'Bereich' ...	ist ein/eine _____
'fliegt mit' ...	ist ein/eine _____
'Vorname' ...	ist ein/eine _____
'Sitzplatz' ...	ist ein/eine _____
'besitzt' ...	ist ein/eine _____
'Passagier' ...	ist ein/eine _____

**Komponententypen:** (starke) Entität (entity), schwache Entität (weak entity), identifizierende Beziehung (identifying relation), unäre Beziehung (unary relation), binäre Beziehung (binary relation), ternäre Beziehung (ternary relation), Generalisierung (generalization), Attribut (attribute), Schlüsselattribut (key attribute), mehrwertiges Attribut (multi-valued attribute), zusammengesetztes Attribut (composed attribute), abgeleitetes Attribut (derived attribute).

- b. (5 Punkte) Führen Sie das ER Diagramm in ein relationales Schema über. Geben Sie pro Relation auch explizit den Primärschlüssel bzw. vorhandene Fremdschlüsselbeziehungen mittels  $\diamond$ -Notation an.
- c. (5 Punkte) Führen Sie Ihr relationales Schema aus Aufgabe b) in ein physisches Schema über. Erstellen Sie dazu mit Hilfe der DDL (Data Definition Language) die benötigten Tabellen (inkl. Primär- und Fremdschlüssel) und geben Sie die entsprechenden CREATE-Anweisungen an. Wählen Sie die Datentypen entsprechend der zu speichernden Information aus.

#### Aufgabe 5 [Grundlagen von Datenbanksystemen: 10 Punkte]

Beschreiben Sie die **drei** typischen Phasen des Datenbankentwurfsprozesses, welche auf eine natürlichsprachige Anforderungsbeschreibung eines zu erstellenden Datenbanksystems folgen. Geben Sie je Phase auch die verwendeten Werkzeuge und/oder aus der Phase resultierenden Ergebnisse (Artefakte) an.