

Datenbanksysteme (2h)	schriftliche Einzelprüfung	21.04.2010		1
-----------------------	----------------------------	------------	--	---

Aufgabe 1 [Relationale Abfragen: 30 Punkte]

Gegeben ist folgendes vereinfachtes Relationenschema:

flug (fnr, ftyp, abf_ort, abf_datum, ank_ort, ank_datum) FK: flug.ftyp \diamond flugzeug

flugzeug (typ, hname, sitzplaetze, leistung, geschwindigkeit) FK: flugzeug.hname \diamond hersteller

hersteller (name, homepage, umsatz)

flugzeug.typ IN {'Airbus A320', 'Boeing 747', 'Messerschmitt Me 262', 'C-130 Hercules'}

hersteller.name IN {'Airbus', 'Boing', 'Junkers', 'Lockheed'}

leistung ... in PS, geschwindigkeit ... in km/h, umsatz ... in EUR pro Jahr

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

- (3 Punkte) Ermitteln Sie Flugzeugtyp, Leistung und Geschwindigkeit aller Flugzeuge die von 'Airbus' oder 'Boeing' stammen und zwischen 300 und 400 (jeweils inklusive) Sitzplätze verfügen.
- (4 Punkte) Ermitteln Sie Name und Homepage des umsatzstärksten Herstellers.
- (5 Punkte) Nennen Sie alle Paare von Flügen, die die gleiche Sitzplatzanzahl, den gleichen Abflugort und den gleichen Ankunftszeitpunkt besitzen und geben Sie die beiden Flugnummern an.

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

- (3 Punkte) Ermitteln Sie Flugzeugtyp, Leistung und Geschwindigkeit aller Flugzeuge die von 'Airbus' oder 'Boeing' stammen und zwischen 300 und 400 (jeweils inklusive) Sitzplätze verfügen.
- (4 Punkte) Ermitteln Sie Name und Homepage des umsatzstärksten Herstellers.
- (5 Punkte) Ermitteln Sie je Herstellername die minimale, durchschnittliche und maximale Sitzplatzanzahl der gebauten Flugzeuge und ordnen Sie diese Informationen nach dem Umsatz des Herstellers (absteigend).
- (6 Punkte) Nennen Sie Homepage und Gesamtsitzplatzanzahl aller Flüge die mit einem Flugzeug von 'Lockheed' vom Flughafen 'Wien' aus starten.

Aufgabe 2 [Query Optimierung: 30 Punkte]

Gegeben ist folgende Abfrage in relationaler Algebra, basierend auf den beiden Relationenschemata $R_1(I, N, T, E, R)$ und $R_2(S, K, Y)$:

$$\pi_{I,R,S,K}(\pi_{I,N,T,R,S,K,Y}(\sigma_{(I>=1 \wedge T<9 \wedge E=7 \wedge Y<=5)}(R_1 \bowtie_{(N=S)} R_2)))$$

- (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 2500 und die Relation R_2 2000 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 $\frac{1}{5}$ (Annahme der Unabhängigkeit).

Geben Sie die Kosten (Anzahl von Blockzugriffen) für

- den ursprünglichen Ausdruck und
- den umgeformten Ausdruck an.

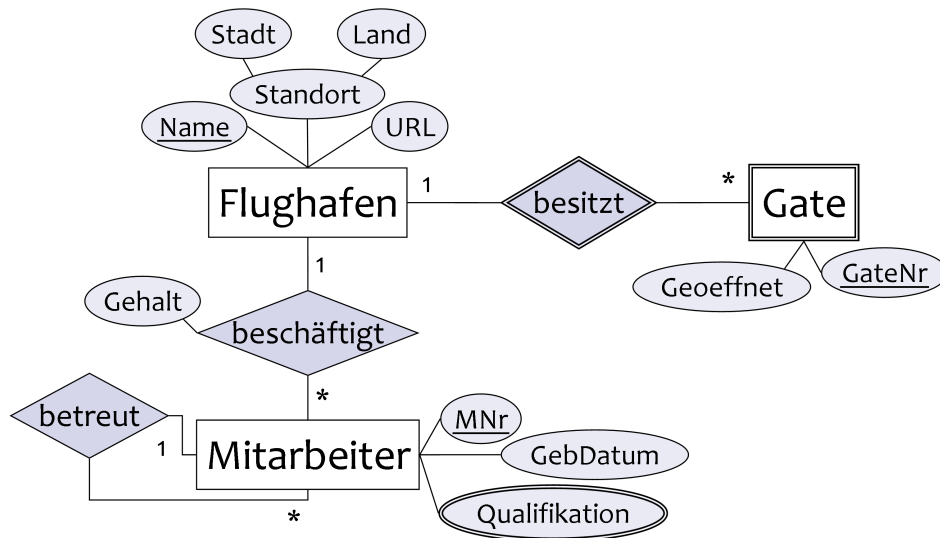
Nehmen Sie an, dass die Abarbeitung der Ausdrücke Pipelining nutzt.

Aufgabe 3 [Formaler Datenbankentwurf: 15 Punkte]

Gegeben ist folgende funktionale Abhängigkeit:

$$RS = (\{T, Y, R, O, L, E, A, N\}, \{RL \rightarrow E, OYR \rightarrow LE, L \rightarrow Y, N \rightarrow R, O \rightarrow NL, T \rightarrow YR, LY \rightarrow EN\})$$

- (5 Punkte) Geben Sie für RS die minimale Überdeckung der funktionalen Abhängigkeiten an.
- (5 Punkte) Bestimmen Sie für RS alle Schlüsselkandidaten.



c. (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:

- | | |
|---------------------|--------------------|
| 'Gate' ... | ist ein/eine _____ |
| 'betreut' ... | ist ein/eine _____ |
| 'Standort' ... | ist ein/eine _____ |
| 'MNR' ... | ist ein/eine _____ |
| 'besitzt' ... | ist ein/eine _____ |
| 'Land' ... | ist ein/eine _____ |
| 'beschäftigt' ... | ist ein/eine _____ |
| 'Qualifikation' ... | ist ein/eine _____ |
| 'Gehalt' ... | ist ein/eine _____ |
| 'Flughafen' ... | 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 [Speicher-, File- und Indexstrukturen: 10 Punkte]

- a. (5 Punkte) Was versteht man unter der 'Byte-String Repräsentation'? Welche Probleme ergeben sich bei der Organisation von Daten daraus, und wie kann man diese lösen?

Datenbanksysteme (2h)	schriftliche Einzelpruefung	21.04.2010		3
-----------------------	--------------------------------	------------	--	---

- b. (5 Punkte) Was versteht man unter der 'Zeiger Methode'? Welche Probleme ergeben sich bei der Organisation von Daten daraus, und wie kann man diese lösen?