

Datenbanksysteme (2h)	schriftliche Einzelprüfung	26.01.2010	1
-----------------------	----------------------------	------------	---

Aufgabe 1 [Relationale Abfragen: 30 Punkte]

Gegeben ist folgendes Relationenschema:

spieler (svnr, vorname, nachname, position, gebjahr)
verein (name, stadt, land, gruendungsjahr)
vertrag (name, svnr, datum, dauer, gehalt)

verein.name \diamond *vertrag*, *spieler.svnr* \diamond *vertrag*
spieler.position IN {'Tor', 'Verteidigung', 'Mittelfeld', 'Sturm'}
dauer ... in Jahren
gehalt ... (monatliches Gehalt) in Euro

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

- (3 Punkte) Ermitteln Sie Name und Geburtsjahr aller Feldspieler (Spieler die nicht im 'Tor' spielen) die zwischen 1985 und 1990 geboren sind.
- (4 Punkte) Ermitteln Sie Name und Position des/der jüngsten Spieler/s.
- (5 Punkte) Ermitteln Sie die Namen aller Spieler, die bei zumindest zwei verschiedenen Vereinen Verträge mit einer Laufzeit (dauer) von jeweils maximal zwei Jahren hatten.

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

- (3 Punkte) Ermitteln Sie Name und Geburtsjahr aller Feldspieler (Spieler die nicht im 'Tor' spielen) die zwischen 1985 und 1990 geboren sind.
- (4 Punkte) Ermitteln Sie Name und Position des/der jüngsten Spieler/s.
- (5 Punkte) Ermitteln Sie die durchschnittliche Jahresgehaltssumme aller Spieler, die am 01.01.2010 ('01-JAN-10') einen Vertrag bei einem Verein in 'Salzburg' (in 'Österreich') unterschrieben haben. Diese berechnet sich aus der Summe der Spielergelälter dividiert durch die Laufzeit (dauer) der Verträge.
- (6 Punkte) Ermitteln Sie die Namen aller Spieler, die sowohl bei einem Verein in 'Deutschland' als auch in 'Österreich', jedoch noch nie in einem Verein in 'Spanien' vertraglich angestellt waren.

Aufgabe 2 [Query Optimierung: 30 Punkte]

Gegeben ist folgende Abfrage in relationaler Algebra, basierend auf den beiden Relationenschemata $R_1(R, E, D, B, U, L)$ und $R_2(S, A, Z, G)$:

$$\pi_{R,A,U,S}(\pi_{R,E,U,S,A,G}(\sigma_{(E>2 \wedge D>5 \wedge Z<10)}(R_1 \bowtie_{(B=A)} R_2)))$$

- (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 1000 und die Relation R_2 5000 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 $\frac{1}{10}$ (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: 15 Punkte]

Gegeben ist folgende funktionale Abhängigkeit:

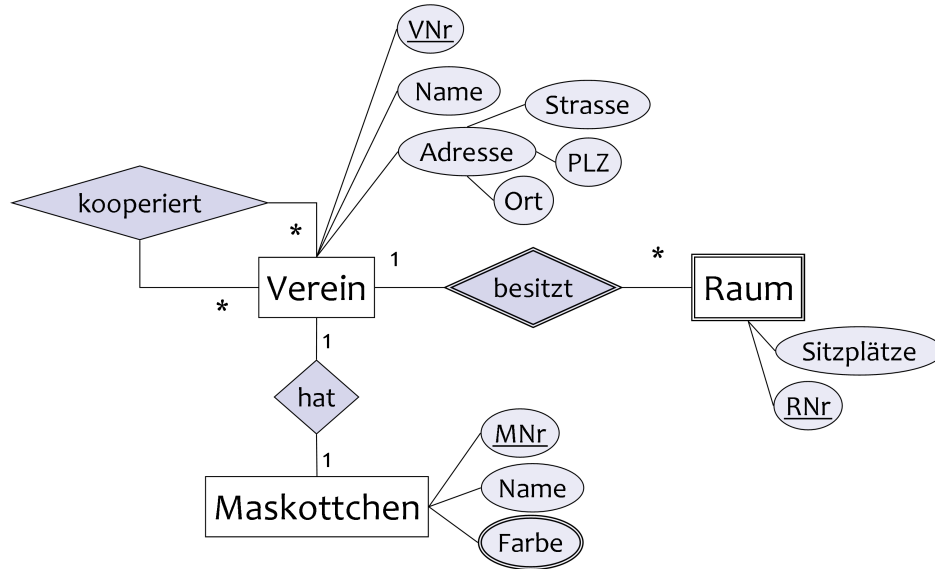
$$RS = (\{R, E, D, B, U, L, S, A, Z, G\}, \{R \rightarrow ED, U \rightarrow E, BED \rightarrow UL, A \rightarrow D, SA \rightarrow G, UE \rightarrow AL, DU \rightarrow GL, B \rightarrow AU\})$$

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

b. (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:

- | | | |
|-------------------|--------------|-------|
| 'Raum' ... | ist ein/eine | _____ |
| 'koopertiert' ... | ist ein/eine | _____ |
| 'Maskottchen' ... | ist ein/eine | _____ |
| 'Adresse' ... | ist ein/eine | _____ |
| 'VNr' ... | ist ein/eine | _____ |
| 'besitzt' ... | ist ein/eine | _____ |
| 'hat' ... | ist ein/eine | _____ |
| 'Name' ... | ist ein/eine | _____ |
| 'Farbe' ... | ist ein/eine | _____ |
| 'Strasse' ... | 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 [Indexstrukturen: 10 Punkte]

Datenbanksysteme (2h)	schriftliche Einzelpruefung	26.01.2010		3
-----------------------	--------------------------------	------------	--	---

- a. (5 Punkte) Erklären Sie die Begriffe (1) **Primärindex**, (2) **Sekundärindex**, (3) **dichter Index**, (4) **dünner Index** und (5) **multilevel Index** jeweils kurz.
- b. (5 Punkte) Zeichnen Sie eine Skizze anhand derer Sie den (6) **B+-Baum Index** und dessen Struktur und Anordnung der einzelnen Knoten erläutern. Des Weiteren nennen Sie zumindest einen Vor- und einen Nachteil von B+-Baum Indizes gegenüber der Verwendung von Index-sequentiellen Dateien.