

Datenbanksysteme (DBS VO)	schriftliche Einzelprüfung	19.04.2013		1
------------------------------	-------------------------------	------------	--	---

Aufgabe 1 [Relationale Abfragen: 30 Punkte]

Gegeben ist folgendes vereinfachtes Relationenschema einer Marathon-Verwaltung:

```

athlet (id, vorname, zuname, geschlecht, geburtsjahr)
PK: id
bewerb (id, jahr, ort, art)
PK: id
teilnahme (athlet_id, bewerb_id, platzierung, stunden, minuten, sekunden)
PK: athlete_id, bewerb_id
FK: athlete_id ◊ athlete
FK: bewerb_id ◊ bewerb

athlet.geschlecht IN {"MÄNNLICH", "WEIBLICH"}
bewerb.art IN {"VOLL", "HALB", "STAFFEL"}

```

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

- (3 Punkte) Ermitteln Sie die id's aller AthletInnen, die jemals eine Platzierung unter den ersten drei erreicht haben.
- (4 Punkte) Ermitteln Sie Vor- und Zunamen der ältesten AthletInnen (mit minimalem Geburtsjahr).
- (5 Punkte) Ermitteln Sie Zuname und Geburtsjahr aller AthletInnen, die niemals an einer Staffel teilgenommen haben.

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

- (3 Punkte) Ermitteln Sie die id's aller AthletInnen, die jemals eine Platzierung unter den ersten drei erreicht haben.
- (4 Punkte) Ermitteln Sie Vor- und Zunamen der ältesten AthletInnen (mit minimalem Geburtsjahr).
- (5 Punkte) Ermitteln Sie für AthletInnen, die an mehr als 3 Bewerbungen teilgenommen haben, den Zunamen und die jeweils schlechteste jemals erreichte Platzierung.
- (6 Punkte) Ermitteln Sie die Zunamen aller AthletInnen, die niemals an einem Bewerb in Wien teilgenommen haben, aber zumindest einmal einen ersten Platz erreicht haben.

Aufgabe 2 [Query Optimierung: 30 Punkte]

Gegeben ist folgende Abfrage in relationaler Algebra, basierend auf den beiden Relationenschemata $R_1(P, A, V, O)$ und $R_2(N, U, R, M, I)$:

$$\pi_{A,M}(\pi_{A,V,R,M}(\sigma_{(O < 100 \wedge U > V \wedge I > 42)}(R_1 \bowtie_{(N=P)} R_2)))$$

- (4 Punkte) Stellen Sie den Ausdruck grafisch dar.
- (8 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.
- (8 Punkte) Nehmen Sie an, dass die Relation R_1 5000 und die Relation R_2 12000 Datensätze enthält, wobei die Blockgröße für beide Relationen 10 ist.
Für den Join wird das Nested-Loop Verfahren verwendet (Memorygröße 1 Block pro Relation) und die Selektivität aller Selektionen ist jeweils 1/10 (Annahme der Unabhängigkeit).
Geben Sie die Kosten (Anzahl von Blockzugriffen) für den **optimierten Ausdruck** (aus Punkt b) an.
Nehmen Sie an, dass die Abarbeitung des Ausdrucks Pipelining nutzt.
- (10 Punkte) Schätzen Sie die Kosten für den **optimierten Ausdruck** auch für folgende Fälle:
 - Nested-Loop mit genügend verfügbarem Memory um die Relation R_2 komplett im Speicher zu halten
 - Block-Nested-Loop mit nur einem Block Memory pro Relation
 - Block-Nested-Loop mit genügend verfügbarem Memory, um die Relation R_2 komplett im Speicher zu halten

Datenbanksysteme (DBS VO)	schriftliche Einzelprüfung	19.04.2013		2
------------------------------	-------------------------------	------------	--	---

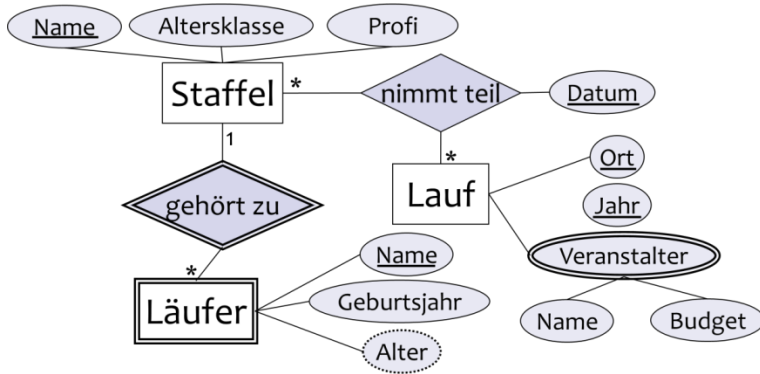
Aufgabe 3 [Formaler Datenbankentwurf: 15 Punkte]

Gegeben ist folgendes Relationenschema mit funktionalen Abhängigkeiten:

$$RS = (\{P, A, V, O, N, U, R, M, I\}, \{N \rightarrow PAV, OAV \rightarrow U, AN \rightarrow OU, VN \rightarrow A, O \rightarrow N\})$$

- (5 Punkte) Geben Sie für RS eine 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

'nimmt teil'	ist ein/eine	_____
'Alter'	ist ein/eine	_____
'Veranstalter'	ist ein/eine	_____
'Staffel'	ist ein/eine	_____
'Datum'	ist ein/eine	_____
'gehört zu'	ist ein/eine	_____
'Läufer'	ist ein/eine	_____
'Geburtsjahr'	ist ein/eine	_____
'Budget'	ist ein/eine	_____
'Lauf'	ist ein/eine	_____

Komponententypen: (starke) Entität (entity), schwache Entität (weak entity), identifizierende Beziehung (identifying relation), reflexive Beziehung (reflexive relation), binäre Beziehung (binary relation), ternäre Beziehung (ternary relation), Generalisierung (generalization), Attribut (attribute), Schlüsselattribut (key attribute), mehrwertiges Attribut (multi-valued attribute), zusammengesetztes Attribut (composite 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 SQL-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.

Datenbanksysteme (DBS VO)	schriftliche Einzelprüfung	19.04.2013		4
------------------------------	-------------------------------	------------	--	---

Aufgabe 5 [Begriffsbestimmungen: 10 Punkte]

Definieren Sie die Begriffe (1) **Oberschlüssel**, (2) **Schlüsselkandidat**, (3) **Schlüssel**, (4) **Primärindex** und (5) **Sekundärindex**.