

Datenbanksysteme (2h)	schriftliche Einzelprüfung	20.04.2012	1
-----------------------	----------------------------	------------	---

Aufgabe 1 [Relationale Abfragen: 30 Punkte]

Gegeben ist folgendes vereinfachtes Relationenschema einer Architektur-Datenbank:

bauwerk (id, name, ort, besucheranzahl)

PK: id

architekt (id, name, geburtsjahr, todesjahr, nationalitaet)

PK: id

erstellt (id_bauwerk, id_architekt, beginn_jahr, ende_jahr, kosten)

PK: id_bauwerk, id_architekt

FK: id_bauwerk \diamond bauwerk

FK: id_architekt \diamond architekt

architekt.nationalitaet IN {"BRD", "USA", "AUT"}

erstellt.kosten ... in Euro

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

- (3 Punkte) Ermitteln Sie Name und Geburtsjahr aller ArchitektInnen, die einer der Nationalitäten BRD oder AUT angehören und deren Todesjahr zwischen 1960 und 1970 (jeweils inklusive) liegt.
- (4 Punkte) Ermitteln Sie die Namen der Bauwerke mit der geringsten Anzahl an Besuchern.
- (5 Punkte) Ermitteln Sie Name und Ort der Bauwerke, die nicht von einem Architekten aus Österreich (AUT) errichtet wurden.

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

- (3 Punkte) Ermitteln Sie Name und Geburtsjahr aller ArchitektInnen, die einer der Nationalitäten BRD oder AUT angehören und deren Todesjahr zwischen 1960 und 1970 (jeweils inklusive) liegt.
- (4 Punkte) Ermitteln Sie die Namen der Bauwerke mit der geringsten Anzahl an Besuchern.
- (5 Punkte) Ermitteln Sie für ArchitektInnen, die an der Errichtung von mindestens 5 Bauwerken beteiligt waren, den Namen (des/der Architekten/in), sowie die durchschnittliche Besucheranzahl aller von ihm/ihr errichteten Bauwerke.
- (6 Punkte) Ermitteln Sie Name und Ort der Bauwerke, an deren Errichtung kein/e ArchitektIn der Nationalität USA beteiligt war, aber zumindest ein/e ArchitektIn, die vor 1960 geboren wurde.

Aufgabe 2 [Query Optimierung: 30 Punkte]

Gegeben seien drei Relationen $R_1(U, X)$, $R_2(V, X)$ und $R_3(W, X)$ sowie die folgende Abfrage:

$$\pi_{X,U}((R_1 \bowtie R_2 \bowtie R_3))$$

- (8 Punkte) Finden Sie drei mögliche, unterschiedliche Abarbeitungsreihenfolgen und stellen Sie diese grafisch in Form von Expression-Trees dar.
- (22 Punkte) Ermitteln Sie die beste der drei Abarbeitungsreihenfolgen, indem Sie für jede den Aufwand abschätzen. Nehmen Sie dazu an, dass R_1 3000, R_2 2000 und R_3 4000 Datensätze enthält, wobei die Blockgröße für R_1 und R_2 50, für R_3 20 ist.
Für alle Joins wird das Nested-Loop Verfahren verwendet (Memorygröße 1 Block pro Relation) und die Selektivität der Joins ist jeweils $\frac{1}{2}$ (Annahme der Unabhängigkeit).
Sie dürfen davon ausgehen, dass die Abarbeitung der Ausdrücke Pipelining nutzt.

Aufgabe 3 [Formaler Datenbankentwurf: 15 Punkte]

Gegeben ist folgendes Relationenschema mit funktionalen Abhängigkeiten:

$$RS = (\{O, P, E, R, N, H, A, U, S\}, \{R \rightarrow N, EPR \rightarrow H, N \rightarrow OEP, NP \rightarrow HR, EN \rightarrow P\})$$

- (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]

