

Datenbanksysteme (DBS VO)	schriftliche Einzelprüfung	08.03.2013		1
------------------------------	-------------------------------	------------	--	---

Aufgabe 1 [Relationale Abfragen: 30 Punkte]

Gegeben ist folgendes vereinfachtes Relationenschema einer Film-Datenbank:

```

film (titel, erscheinungsjahr, genre, budget)
PK: titel, erscheinungsjahr
schauspieler (id, vorname, zuname, geburtsjahr)
PK: id
oscar (kategorie, jahr, id, titel, erscheinungsjahr)
PK: kategorie, jahr
FK: id ∅ schauspieler
FK: titel, erscheinungsjahr ∅ film

film.genre IN {"ACTION", "KOMÖDIE", "DRAMA"}

```

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

- (3 Punkte) Ermitteln Sie die Titel und Erscheinungsjahre aller Filme, für die ein Oscar in der Kategorie „Regie“ zwischen 1940 und 1950 vergeben wurde.
- (4 Punkte) Ermitteln Sie die Titel der Filme, die mit den höchsten Kosten gedreht wurden (mit maximalem Budget).
- (5 Punkte) Ermitteln Sie Zuname und Geburtsjahr aller SchauspielerInnen, die niemals einen Oscar erhalten haben.

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

- (3 Punkte) Ermitteln Sie die Titel und Erscheinungsjahre aller Filme, für die ein Oscar in der Kategorie „Regie“ zwischen 1940 und 1950 vergeben wurde.
- (4 Punkte) Ermitteln Sie die Titel der Filme, die mit den höchsten Kosten gedreht wurden (mit maximalem Budget).
- (5 Punkte) Ermitteln Sie für SchauspielerInnen, die mehr als 4 Oscars erhalten haben, den Zunamen und das durchschnittliche Budget aller Filme, für die der/die SchauspielerIn Oscars erhalten hat.
- (6 Punkte) Ermitteln Sie die Oscar-Kategorien, für die niemals ein Oscar an einen Film aus dem Genre „DRAMA“ vergeben wurde, aber zumindest einmal an einen/eine SchauspielerIn mit einem Geburtsjahr nach 1980.

Aufgabe 2 [Query Optimierung: 30 Punkte]

Gegeben ist folgende Abfrage in relationaler Algebra, basierend auf den beiden Relationenschemata $R_1(J, A, M, E, S)$ und $R_2(B, O, N, D)$:

$$\pi_{S,O}(\pi_{S,A,D,O}(\sigma_{(A < 100 \wedge M > E \wedge D > 42)}(R_1 \bowtie_{(A=D)} R_2)))$$

- (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 8000 und die Relation R_2 2000 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	08.03.2013		2
------------------------------	-------------------------------	------------	--	---

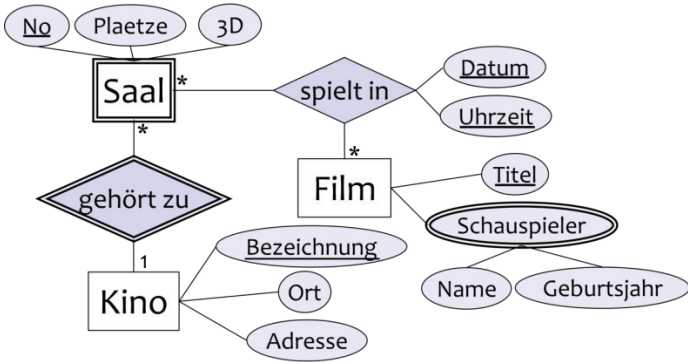
Aufgabe 3 [Formaler Datenbankentwurf: 15 Punkte]

Gegeben ist folgendes Relationenschema mit funktionalen Abhängigkeiten:

$$RS = (\{J, A, M, E, S, B, O, N, D\}, \{S \rightarrow JAM, EAM \rightarrow B, AS \rightarrow EB, MS \rightarrow A, E \rightarrow S\})$$

- (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

- | | | |
|----------------|--------------|-------|
| 'Schauspieler' | ist ein/eine | _____ |
| 'Ort' | ist ein/eine | _____ |
| 'Titel' | ist ein/eine | _____ |
| 'Saal' | ist ein/eine | _____ |
| 'Kino' | ist ein/eine | _____ |
| 'spielt in' | ist ein/eine | _____ |
| 'gehört zu' | ist ein/eine | _____ |
| 'Geburtsjahr' | ist ein/eine | _____ |
| 'Film' | ist ein/eine | _____ |
| 'Plaetze' | 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	08.03.2013		4
------------------------------	-------------------------------	------------	--	---

Aufgabe 5 [Begriffsbestimmungen: 10 Punkte]

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