

1 Aufgabe 1: ER-Modellierung (13 Punkte)

Gegeben ist das in Abbildung 1 dargestellte ER-Modell. Die Schlüsselattribute sind durch Unterstreichen gekennzeichnet. Die Kardinalitäten sind als *Teilnehmerkardinalitäten* angegeben.

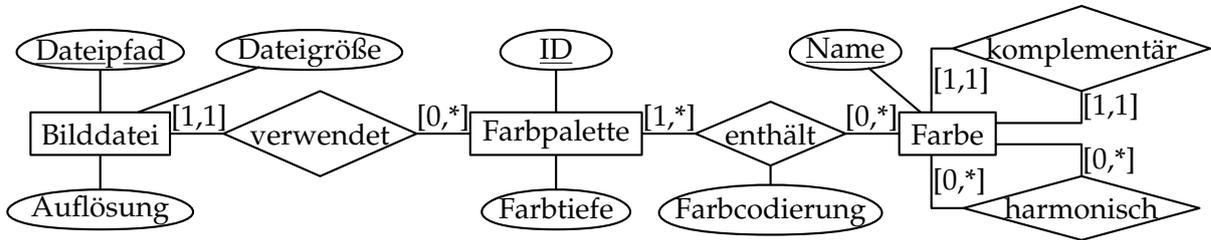


Abbildung 1: Gegebenes ER-Modell

a) Beantworten Sie die folgenden Fragen. Für jede richtige Antwort bekommen Sie 1/2 Punkt. Falsche und nicht beantwortete Fragen werden mit 0 Punkten bewertet. Bitte beantworten Sie die Fragen kurz und präzise; Stichpunkte genügen.

(3 Punkte)

Fragen zu dem in Abbildung 1 dargestellten ER-Modell:

- (i) Mit wievielen Entitäts-Typen steht *Farbe* in Beziehung?
- (ii) Stellen Sie die *verwendet*-Beziehung semantikerhaltend in Standardkardinalität dar.

Allgemeine Fragen zu ER- und EER-Modellen:

- (iii) Welche Eigenschaften gelten für einen Schlüsselkandidaten?
- (iv) Bei welchen Konstrukten können Teilnehmerkardinalitäten und Standardkardinalitäten unterschiedliche Ausdrucksmöglichkeiten haben?
- (v) Wodurch wird die IST-Beziehung im EER-Modell ersetzt?
- (vi) Wie werden Primärschlüssel im EER-Modell gekennzeichnet?

b) Überführen Sie das ER-Modell aus Abbildung 1 kapazitätserhaltend in ein relationales Modell. Optimieren Sie die Relationen dabei soweit wie möglich mit den in der Vorlesung vorgestellten Methoden. Bitte geben Sie nur die Attribute an, nicht aber deren Datentyp. Geben Sie die Primärschlüssel durch Unterstreichen und Fremdschlüssel durch Pfeile an. Verwenden Sie die auf dem Antwortblatt vorgegebene Struktur.

Hinweis: Es müssen nicht alle Felder ausgefüllt werden.

(5 Punkte)

c) Modellieren Sie folgenden Sachverhalt durch ein EER-Modell:

- Sie können Entitäten vom Typ *Massivhaus* erfassen. Jedes *Massivhaus* hat ein Attribut *Wandmaterial*.
- Sie können auch Entitäten vom Typ *Fachwerkhaus* speichern. Jedes *Fachwerkhaus* hat ein Attribut *Holzart*.
- Einige der *Massivhäuser* und *Fachwerkhäuser* sind *Wohnhäuser*.
- Modellieren Sie den Typ *Wohnhaus* unter der Annahme, dass es auch *Wohnhäuser* gibt, die weder *Fachwerkhaus* noch *Massivhaus* sind.
- Jedes *Wohnhaus* hat ein Attribut *Anzahl_Zimmer*.
- Für eine genauere Erfassung werden *Wohnhäuser* in *Einfamilienhäuser* und *Mehrfamilienhäuser* unterschieden. Jedes *Wohnhaus* ist entweder ein *Einfamilienhaus* oder ein *Mehrfamilienhaus*.
- Jedes *Einfamilienhaus* hat ein Attribut *Einzugsdatum*.
- Jedes *Mehrfamilienhaus* hat ein Attribut *Anzahl_Wohnungen*.

(5 Punkte)

Aufgabe 2 – Histories (15 Punkte)

a) Zeichnen Sie den Verlauf der Sperrenbelegung für das schwache und für das strikte Zwei-Phasen-Sperrprotokoll (2PL) in die Koordinatensysteme auf dem Antwortblatt ein. Kennzeichnen Sie jeweils die beiden Phasen des 2PL. Wie unterscheiden sich die beiden Sperrverfahren? (2.5 Punkte)

b) (i) Gegeben sei folgender Schedule S mit den Transaktionen T_1, T_2 und T_3 :

$$S = w_2[x] \ w_3[y] \ w_1[x] \ r_3[y] \ r_1[y] \ r_2[x] \ r_2[y] \ c_1 \ c_2 \ c_3$$

(1) Ergänzen Sie den Schedule um die nötigen Aktionen zum Setzen (z.B. $rl_1[a]$ oder $wl_1[a]$) und Rücksetzen (z.B. $ul_1[a]$) von Sperren unter Anwendung des strikten 2PL. Das Setzen von Sperren soll dabei zum spätest möglichen Zeitpunkt erfolgen, wobei das Rücksetzen von Sperren keinen zeitlichen Aufschub verursacht.

(2) Ist der Schedule ausführbar?

(3) Geben Sie alle möglichen Schedules $S' \neq S$ an, die sich ausschließlich dadurch unterscheiden, dass nicht das strikte 2PL sondern das schwache 2PL verwendet wird.

(4) Gibt es einen solchen Schedule S' , sodass dieser ausführbar ist?

(ii) Gegeben sei folgende Historie H_1 mit den Transaktionen T_4 und T_5 :

$$H_1 = r_5[y] \ w_4[z] \ r_4[x] \ w_4[x] \ a_4 \ w_5[y] \ c_5$$

Erweitern Sie H_1 zu H_1' , indem Sie die Operationen von

$$T_6 = r_6[z] \ w_6[x] \ c_6$$

so hinzufügen, dass bei Ausführung ein Deadlock entsteht.

(iii) Sind H_1 und H_1' konfliktäquivalent? Begründen Sie Ihre Antwort.

(5.5 Punkte)

c) (i) Geben Sie die Definition der Rücksetzbarkeitsklassen RC und ACA an.

(ii) Welche zusätzliche Eigenschaft erfordert die Rücksetzbarkeitsklasse STRICT?

(iii) Es sei folgende unvollständige Historie H_2 mit den Transaktionen T_7, T_8 und T_9 gegeben:

$$H_2 = w_9[y] \ w_7[x] \ r_9[y] \ r_7[y] \ w_8[x] \ r_8[x] \ r_8[y]$$

Es sei weiter H_3 die Historie, die H_2 um die Operationen c_1, c_2 und c_3 ergänzt.

(1) Geben Sie die notwendigen Bedingungen an die Reihenfolge der in H_3 ergänzten Operationen an, damit H_3 RC erfüllt aber nicht ACA. Geben Sie zudem eine Beispielhistorie H_3 an, für die dies gilt.

(2) Geben Sie die notwendigen Bedingungen an die Reihenfolge der Operationen in H_3 an, damit H_3 ACA erfüllt aber nicht STRICT. Geben Sie zudem eine Beispielhistorie H_3 an, für die dies gilt.

(3) Wie viele Möglichkeiten gibt es c_1, c_2 und c_3 anzuordnen, damit H_3 STRICT erfüllt? Geben Sie zudem eine Beispielhistorie H_3 an, für die dies gilt.

(7 Punkte)

Aufgabe 3 – SQL (15 Punkte)

Wir wollen untersuchen, wie gut sich aktuelle Modelle von Google zur Textanalyse eignen, und diese dazu relational speichern. Ein Modell weist jedem Wort eines Textkorpus (z. B. Wikipedia) einen 3-dimensionalen Vektor (x, y, z) zu. Dazu kommen weitere Eigenschaften wie Typ (z. B. Verb) und Themenbereich (z. B. Musik). Interessant hierbei ist, dass ähnliche Wörter generell kürzere Distanzen zueinander haben. Daher ist z. B. der Abstand zwischen den Vektoren von 'Klavier' und 'Gitarre' kleiner als der zwischen 'Klavier' und 'Tasse'.

Erläuterung: Nehmen Sie die nachfolgend abgebildete Tabellenstruktur mit sinnvollen Datentypen und Schlüssel an. Die Daten in den Tabellen sind exemplarisch, geben Sie daher für die unten gestellten Aufgaben immer generelle Lösungen an! Alle Aufgaben lassen sich mit SQL-Konstrukten lösen, die in der Vorlesung vorgestellt wurden. Die Teilaufgaben sind voneinander unabhängig.

- a) Formulieren Sie eine Anfrage, die bestimmt, zu welchem Bereich das Wort 'Fußball' gehört. Geben Sie dessen Namen aus. (2 Punkte)
- b) Geben Sie alle Substantive des Bereichs 'Musik' aus, die mit der Buchstabenfolge 'AC' beginnen. (3 Punkte)
- c) Bestimmen Sie diejenigen Worttypen, denen mindestens 1000 Wörter zugeordnet sind. Geben Sie die jeweiligen Typ-Namen und die Anzahl der zugeordneten Wörter aus. (2 Punkte)
- d) Erstellen Sie eine Sicht $Abstand(w_1, w_2, abstand)$. Sie besteht aus allen Wortpaaren $(w_1.text, w_2.text)$ und dem zugehörigen Abstand. Beachten Sie: die Abstände (w_1, w_2) und (w_2, w_1) sind identisch. Daher soll jedes Wortpaar **nur einmal** in der Sicht enthalten sein, unabhängig von der Reihenfolge der Wörter, und es soll $w_1 \neq w_2$ gelten. Der Abstand zweier Vektoren wird wie folgt berechnet:
$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$
Hinweis: Sie können die Funktionen $SQR(x)=x^2$ und $SQRT(x)=\sqrt{x}$ benutzen. (3 Punkte)
- e) Betrachten Sie die Sicht $Abstand$ aus Aufgabenteil d) als gegeben. Bestimmen Sie damit die drei kleinsten existierenden Abstände zum Wort 'Klavier'. (2 Punkte)
- f) Betrachten Sie die Sicht $Abstand$ aus Aufgabenteil d) als gegeben. Geben Sie für das Wort 'Mozart' alle Wörter und den Namen ihres zugehörigen Bereichs aus, deren Abstand geringer als 0.20 ist. (3 Punkte)

Wort						
id	text	typ	bereich	x	y	z
1	Klavier	1	1	0.1	0.1	0.1
2	Gitarre	1	1	0.1	0.2	0.1
3	Mozart	1	1	0.1	0.1	0.2
4	ACDC	1	1	0.1	0.2	0.2
5	komponieren	2	1	0.2	0.1	0.2
6	Tasse	1	4	0.9	0.3	0.1
7	Fußball	1	5	0.1	0.7	0.9

Typ		
id	name	kurzform
1	Substantiv	N
2	Verb	V
3	Adjektiv	AD
4	Pronomen	P
5	Artikel	AR

Bereich	
id	name
1	Musik
2	Politik
3	Wirtschaft
4	Haushalt
5	Sport

Aufgabe 4 – Funktionale Abhängigkeiten (17 Punkte)

- a) Gegeben sei die Relation $R(A, B, C, D, E, F, G, H, I, J)$ mit der folgenden Menge an funktionalen Abhängigkeiten:

$$F = \{ \begin{array}{l} A B \rightarrow C \\ A \rightarrow D E \\ B \rightarrow F \\ F \rightarrow G H \\ D \rightarrow I J \end{array} \}$$

- (i) Sind die folgenden Aussagen richtig oder falsch? Für jede Aussage gibt es bei einer korrekten Antwort 1/3 Punkt. Für jede Aussage, für die die Antwort nicht korrekt ist oder fehlt, gibt es 0 Punkte.

Funktionale Abhängigkeit	RICHTIG	FALSCH
$A B \rightarrow D E$	<input type="checkbox"/>	<input type="checkbox"/>
$B \rightarrow C$	<input type="checkbox"/>	<input type="checkbox"/>
$A \rightarrow C$	<input type="checkbox"/>	<input type="checkbox"/>
$A \rightarrow J$	<input type="checkbox"/>	<input type="checkbox"/>
$B \rightarrow G$	<input type="checkbox"/>	<input type="checkbox"/>
$F \rightarrow B$	<input type="checkbox"/>	<input type="checkbox"/>

- (ii) Geben Sie die Menge aller Schlüssel der Relation R an.
 (iii) In welcher höchsten Normalform befindet sich R ? Gehen Sie dabei davon aus, dass alle Attribute atomare Domänen haben. Begründen Sie.
 (iv) Geben Sie eine verbundtreue und abhängigkeitsbewahrende Zerlegung von R an, die in der nächst höheren Normalform als R liegt.
 (v) Ist die Zerlegung von R in $R_1(\underline{A}, \underline{B}, \underline{C})$, $R_2(\underline{A}, \underline{D}, \underline{E})$, $R_3(\underline{D}, \underline{I}, \underline{J})$ und $R_4(\underline{B}, \underline{F})$ verbundtreu? Begründen Sie.

(6 Punkte)

- b) Gegeben sei die Relation $R(A, B, C, D, E, F)$ mit der folgenden Menge an funktionalen Abhängigkeiten:

$$F = \{ \begin{array}{l} A \rightarrow B C \\ C D \rightarrow E \\ B \rightarrow D \\ E \rightarrow A \\ F \rightarrow A \end{array} \}$$

- (i) Ist folgende Zerlegung abhängigkeitsbewahrend? Geben Sie entsprechend einen formalen Nachweis bzw. ein Gegenbeispiel an.

$$R_1(A, B, C)$$

$$R_2(A, D, E, F)$$

- (ii) Ist folgende Zerlegung abhängigkeitsbewahrend? Geben Sie entsprechend einen formalen Nachweis bzw. ein Gegenbeispiel an.

$$R_1(A, B, C)$$

$$R_2(C, D, E, F)$$

(4,5 Punkte)

- c) Gegeben sei die Relation $R(A, B, C, D, E, F, G, H)$ mit der folgenden Menge an funktionalen Abhängigkeiten:

$$F = \{$$

$$C \rightarrow D E$$

$$C G \rightarrow A$$

$$D \rightarrow B H$$

$$D F A \rightarrow G$$

$$D A \rightarrow C$$

$$\}$$

- (i) Berechnen Sie die Abschlüsse der folgenden Attributmengen:

Attributmenge
{C}
{D}
{C, G}
{A, D}
{A, D, F}

- (ii) Welche der folgenden Abhängigkeiten verletzen die BCNF? Begründen Sie jeweils Ihre Antwort.

Abhängigkeit

$$C \rightarrow D E$$

$$C G \rightarrow A$$

$$D \rightarrow B H$$

$$D F A \rightarrow G$$

$$D A \rightarrow C$$

(6,5 Punkte)