



3. Übungsblatt
(Testatwoche: 4. - 6. Mai 2010)

Einführung in Datenbanksysteme
Datenbanken für die Bioinformatik

Heinz Schweppe, Katharina Hahn

Aufgabe 1 (DDL + Constraints)

12 Punkte

Die Tabellen *gruppe* und *benutzer* werden mit folgenden SQL Befehlen erzeugt:

```
CREATE TABLE gruppe(  
  name VARCHAR(15) NOT NULL,  
  gruender VARCHAR(8),  
  gruendungsjahr INT,  
  PRIMARY KEY (name),  
  CONSTRAINT gruendung_ok  
    CHECK (gruendungsjahr > 2000)  
);
```

```
CREATE TABLE benutzer (  
  benutzername VARCHAR(8) NOT NULL,  
  nachname VARCHAR(12) NOT NULL,  
  vorname VARCHAR(10) NOT NULL,  
  geburtstag DATE,  
  haarfarbe VARCHAR(7),  
  familienstand VARCHAR(11),  
  gruppe VARCHAR(15),  
  PRIMARY KEY (benutzername),  
  CONSTRAINT familienstand CHECK (familienstand IN  
    ('ledig', 'verheiratet', 'geschieden', 'verwitwet')),  
  CONSTRAINT fk_group  
    FOREIGN KEY (gruppe) REFERENCES gruppe(name)  
);
```

a) Welche der folgenden Datensätze können in die (anfänglich) leeren Tabellen eingefügt werden? Eingefügte Datensätze werden beim nächsten Schritt als in der Datenbank vorhanden betrachtet. Begründen Sie jeweils Ihre Entscheidung, wenn ein Datensatz nicht eingefügt werden kann.

- ('mir_san_mir','uli_h',2001)
→ gruppe
Datensatz kann eingefügt werden
- ('void', '', 2009)
→ gruppe
Datensatz kann eingefügt werden
- ('kanzlerin', 'angie m.', 2005)
→ gruppe
Datensatz kann eingefügt werden
- ('angie', 'm', 'angela', '1954-07-17', 'wechselnd', 'verheiratet', 'kanzlerin')
→ benutzer
Datensatz kann NICHT eingefügt werden (Haarfarbe zu lang)
- ('tv_addicts','jimmy', 2004)

- gruppe
Datensatz kann eingefügt werden
- ('guido', 'westerwelle', 'guido', '1961-12-27', 'blond', 'ledig', tv_addicts)
→ benutzer
Datensatz kann NICHT eingefügt werden
(tv_addicts ist ohne Anführungszeichen kein VARCHAR)
- ('schnappi', 'das krokodil', 'schnappi', '2004-01-01', 'gruen', 'single', 'tv_addicts')
→ benutzer
Datensatz kann NICHT eingefügt werden
('single' entspricht nicht dem Familienstand-Constraint)
- ('chuckie', 'norris', 'chuck', '1940-03-10', 'grau', 'ledig',
 'ich_krieg_auch_20_prozent_auf_tiernahrung')
Datensatz kann NICHT eingefügt werden
(Die Gruppe ist nicht vorhanden, außerdem wäre der Gruppename zu lang.)
→ benutzer
- ('fischfansBB', 'sharky', 1992)
→ gruppe
Datensatz kann NICHT eingefügt werden (Das Gründungsjahr ist nicht > 2000)
- ('franZZ', 'beckenbauer', 'franz', '1945-09-11', 'grau', 'verheiratet', 'mir_san_mir')
→ benutzer
Datensatz kann eingefügt werden
- ('guido', 'westerwelle', 'guido', '1961-12-27', 'blond', 'ledig', 'void')
→ benutzer
Datensatz kann eingefügt werden
- ('uli_h', 'hoeness', 'uli', '1952-01-15', 'grau', 'verheiratet', 'kanzler')
→ benutzer
Datensatz kann NICHT eingefügt werden (Die Gruppe 'kanzler' existiert nicht.)

b) Auf welche Kardinalitäten (min,max) lassen die Tabellendefinitionen schließen? Begründen Sie ihre Antwort.

gruppe(0,1) ----- (0,*) benutzer

Es kann für eine Gruppe null oder beliebig viele Benutzer geben, ein Benutzer kann Mitglied in höchstens einer Gruppen sein (denn der 'benutzername' ist als Primärschlüssel definiert und kann daher nicht mehrfach eingefügt werden. Aber einem Mitglied muss nicht unbedingt eine Gruppe zugewiesen werden).

c) Wie müssten die Tabellendefinitionen geändert werden, wenn man folgende Forderung umsetzen möchte: *Jede Gruppe besteht aus genau einem Benutzer und Benutzer müssen vor dem 28. April 1986 geboren sein.* Geben Sie die resultierenden Tabellen mit allen Constraints an und beschreiben Sie ihre Vorgehensweise.

Jede Gruppe besteht aus genau einem Benutzer:

Dazu muss man in den Befehl, der die Gruppe erstellt, einen Fremdschlüssel aufnehmen, der nicht NULL sein darf:

```
benutzer VARCHAR(8) NOT NULL,
CONSTRAINT fk_user
FOREIGN KEY (benutzer) REFERENCES benutzer (benutzername)
```

Benutzer müssen vor dem 28. April 1986 geboren sein:

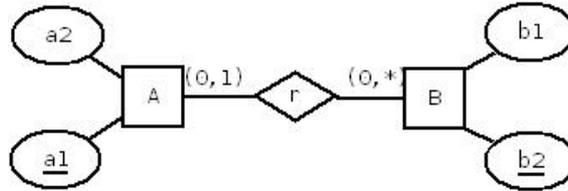
Einfach einen CHECK-Constraint bei Benutzer einfügen:

```
CONSTRAINT geb_ok
CHECK (geburtstag < '1986-04-28')
```

Aufgabe 2 (Funktionale Abhängigkeiten)**6 Punkte**

Betrachten ist das in den Teilaufgaben angegebene ER-Modell, mit den Entitäten A, B und der Relation r.

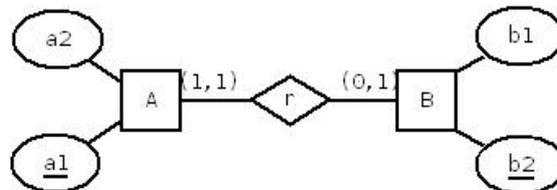
- a) Wie werden die Kardinalitäten (min, max) des folgenden Modells in SQL umgesetzt? Geben Sie jeweils die DDL-Statements an; die Datentypen der Attribute können Sie dabei frei wählen.



```
CREATE TABLE A (
    a1 INT,
    a2 INT,
    PRIMARY KEY(a1)
);

CREATE TABLE B (
    b1 INT,
    b2 INT,
    ra1 INT,
    PRIMARY KEY(b2),
    CONSTRAINT fk_a,
    FOREIGN KEY (ra1) REFERENCES A (a1)
);
```

- b) Wie werden die Kardinalitäten (min,max) des folgenden Modells in SQL umgesetzt? Geben Sie jeweils die DDL-Statements an; die Datentypen der Attribute können Sie dabei frei wählen.



```
CREATE TABLE A (
    a1 INT,
    a2 INT,
    rb2 INT,
    PRIMARY KEY(a1),
    CONSTRAINT fk_b2
    FOREIGN KEY (rb2) REFERENCES B (b2)
);

CREATE TABLE B (
    b1 INT,
    b2 INT,
    ra1 INT NOT NULL,
    PRIMARY KEY(b2),
    CONSTRAINT fk_a1
    FOREIGN KEY (ra1) REFERENCES A (a1),
    UNIQUE (ra1)
);
```

Aufgabe 3 (Funktionale Abhängigkeiten) 6 Punkte

Gegeben ist folgende Relation $R(A, B, C, D)$:

A	B	C	D
1	Tennis	b	a
2	Fußball	ab	b
2	Tennis	ab	b

Welche funktionalen Abhängigkeiten können für das Relationenschema R vorliegen bzw. welche sind ausgeschlossen? „Triviale“ Abhängigkeiten der Form $X \rightarrow Y$, mit $Y \subseteq X$ müssen nicht berücksichtigt werden.

Es gibt insgesamt 18 nicht triviale funktionale Abhängigkeiten auf einzelne Attribute:

$A \rightarrow C$	$A \rightarrow D$	$C \rightarrow A$	$C \rightarrow D$	$D \rightarrow A$	$D \rightarrow C$
$AB \rightarrow C$	$AB \rightarrow D$	$BC \rightarrow A$	$BC \rightarrow D$	$BD \rightarrow A$	$BD \rightarrow C$
$AC \rightarrow D$	$CD \rightarrow A$	$DA \rightarrow C$	$ABC \rightarrow D$	$ABD \rightarrow C$	$BCD \rightarrow A$

Zusätzlich gelten natürlich alle funktionalen Abhängigkeiten auf Attributkombinationen, die sich aus den obigen Abhängigkeiten ableiten lassen, z.B., da $A \rightarrow C$ und $A \rightarrow D$ gilt, gilt auch $A \rightarrow CD$.

Aufgabe 4 (Funktionale Abhängigkeiten)**8 Punkte**

Gegeben ist die Relation $R(A, B, C, D)$

A	B	C	D
a1	b1	c1	d1
a2	b1	c2	d1
a1	b2	c3	d1
a3	b2	c4	d1

mit den folgenden Abhängigkeiten:

1. $B \rightarrow D$
2. $A \rightarrow D$
3. $C \rightarrow ABD$
4. $AB \rightarrow C$

- a) Nennen Sie alle Superschlüssel und Schlüsselkandidaten der Relation. (Ein Superschlüssel ist ein Schlüssel, der nicht unbedingt minimal ist. Die Menge der Superschlüssel ist also eine Obermenge der Kandidatenschlüssel. Ein Superschlüssel ist z.B. R , da $R \rightarrow R$ gilt.)

Superschlüssel:

C, AB, AC, BC, CD, ABC, ABD, ACD, BCD, ABCD
(alle Kombinationen, die AB oder C enthalten)

Schlüsselkandidaten:

AB, C

- b) Die folgenden Zeilen sollen sequentiell in R eingefügt werden. Welche der Zeilen können eingefügt werden, ohne die Abhängigkeiten zu verletzen? Geben Sie die Nummer der Abhängigkeit an, wenn Sie das Einfügen der Zeile verletzt und begründen Sie ihre Antwort.

1. a5, b6, c7, d8 geht
2. a7, b2, c8, d2 geht nicht (1. funktionale Abhängigkeit verletzt)
3. a4, b3, c3, d1 geht nicht (3. funktionale Abhängigkeit verletzt)
4. a2, b1, c5, d1 geht nicht (4. funktionale Abhängigkeit verletzt)

Aufgabe 5 (Normalformen)

12 Punkte

Gegeben sind folgende Relationen mit den funktionalen Abhängigkeiten:

R1(A, D, C, E, F)

F(R1) = {D → E, E → F, ACF → D}

R2(B, C, D)

F(R2) = {BC → D}

- a) Bestimmen Sie alle Schlüsselkandidaten für R1 und R2 und begründen Sie ihre Antwort.

R1: ACD, ACE, ACF (ACF ist ein Schlüsselkandidat und wegen der ersten beiden fkt. Abhängigkeiten kann F durch E oder D im Schlüssel ersetzt werden)

R2: BC (BC → D ist die einzige funktionale Abhängigkeit, die besteht, und damit sind alle Attribute eindeutig durch BC bestimmt.)

- b) In welcher Normalform sind R1 bzw. R2?

R1:

ist in 1. NF (alle Attribute atomar)

ist in 2. NF (Nichtschlüsselattribute vom kompletten Schlüssel abhängig)

ist in 3. NF (keine transitiven Abhängigkeiten von Nichtschlüsselattributen)

Das ist offensichtlich der Fall, da gar keine Nichtschlüsselattribute vorhanden sind

ist NICHT in BCNF (keine transitiven Abhängigkeiten von Schlüsselattributen)

→ die BCNF ist offensichtlich verletzt, es gibt transitive Abhängigkeiten bei den Schlüsselkandidaten

R2:

ist in 1. NF (alle Attribute atomar)

ist in 2. NF (Nichtschlüsselattribute vom kompletten Schlüssel abhängig)

ist in 3. NF (keine transitiven Abhängigkeiten von Nichtschlüsselattributen)

ist in BC-NF (keine transitiven Abhängigkeiten von Schlüsselattributen)

→ R2 ist sogar in BC-NF, denn BC ist Superschlüssel