

- 7) Gegeben sei eine Relation  $R(A, B, C, D, E)$  mit folgenden funktionalen Abhängigkeiten:  
 $A \rightarrow C, AC \rightarrow D, AB \rightarrow E, BCD \rightarrow AE$   
 Welche der folgenden Aussagen sind wahr?  
 a)  $AC \rightarrow D$  ist eine partielle Abhängigkeit.  
 b)  $BCD \rightarrow CD$  ist eine triviale Abhängigkeit.  
 c)  $BCD \rightarrow AE$  ist eine volle Abhängigkeit.  
 d)  $AB$  ist ein Kandidatschlüssel.  
 e)  $AC$  ist ein Superschlüssel.

Für folgende Aufgaben gebe vollständige Lösungen und erkläre die Schritte, die ihr in den Lösungen folgt.

1. (0.5p)  
 a) Definiere die Referenz-Integritätsregel und erläutere die Definition anhand eines Beispiels.  
 b) Definiere eine funktionale Abhängigkeit und erläutere die Definition anhand eines Beispiels.
2. (0.85p) Gegeben sei folgende Relationen Schema:  
 Raum (Id, Name)  
 Vorlesung (Id, Titel, ECTS, ProfessorId)  
 Professor (Id, Name, Vorname, Geburtsdatum)  
 Prüfung (VorlesungId, Datum, RaumId)  
 a) Schreibe eine äquivalente nicht-geschachtelte Abfrage (ohne Unterabfragen) zu der folgenden Abfrage:  

```
SELECT T1.nr1 + T2.nr2 FROM
(SELECT count(*) as nr1 FROM Raum R INNER JOIN Prüfung P ON R.Id = P.RaumId
 INNER JOIN Vorlesung V ON V.Id = P.vorlesungId ) T1,
(SELECT count(*) as nr2 FROM Vorlesung V2 where V2.Id not in (SELECT VorlesungId
FROM Prüfung ) ) T2
```

  
 b) Nehmen wir an, dass folgende Abfragen oft in der gegebenen Datenbank vorkommen:  

```
SELECT * FROM Prüfung WHERE RaumId = ?
SELECT RaumId, Datum FROM Prüfung WHERE RaumId = ? AND Datum > ?
SELECT VorlesungId, Datum FROM Prüfung WHERE VorlesungId = ?
```

Schlage zwei Indexe vor, die ihr erstellen würdet um die Anfragen zu optimieren. Für jeden Index bestimme die Art (geclustert oder nicht-geclustert), die Datenstruktur (Baum oder Hash), die Attributen auf welchen der Index erstellt wird, für welche der gegebenen Anfragen der Index nützlich wäre und ob der Index covering ist oder nicht. Begründe die Antwort!

3. (1.75p) Gegeben sei die Relation  $R(A, B, C, D, E, F, G)$  mit folgenden funktionalen Abhängigkeiten:  
 $B \rightarrow DFG$   
 $BD \rightarrow CE$   
 $AC \rightarrow BF$   
 $DE \rightarrow A$   
 $E \rightarrow B$   
 $G \rightarrow EF$   
 $ADF \rightarrow BC$

- a) Gebe die Definition eines Kandidatschlüssels.  
 b) Finde alle Kandidatschlüssel der Relation R. Begründe.  
 c) Gebe die Definition der dritten Normalform.  
 d) Ist die Relation R in 2NF, 3NF und BCNF? Begründe.

4. (1.75p) Gegeben sei die Relation  $R(A, B, C, D, E, F, G)$  mit folgenden funktionalen Abhängigkeiten:

$$F = \{ ABG \rightarrow E, DG \rightarrow B, B \rightarrow C, G \rightarrow D, C \rightarrow BD, AG \rightarrow BC, B \rightarrow D \}$$

Ist die Relation  $R$  in der dritten Normalform? Begründe. Wenn die Relation nicht in 3NF ist, dann benutze das Synthesealgorithmus um eine 3NF Zerlegung zu finden. Erkläre die Schritte, die ihr folgt. Ist diese Zerlegung verlustlos und abhängigkeitsbewahrend? Begründe.

5. (1p) Gebe ein Beispiel aus der Realität (das auch Sinn macht), unterschiedlich von den Beispielen aus der Vorlesung und Seminar, von einer Relation, die in der zweiten Normalform, aber nicht in der dritten Normalform ist. Erkläre warum diese Relation nicht in der dritten Normalform ist. Erkläre was für Probleme in eurem Schema vorkommen können und gebe ein Beispiel (mit einem Snapshot der Tabelle) wann die beschriebenen Probleme vorkommen. Wie würdet ihr die Tabelle modellieren, sodass die beschriebenen Probleme nicht mehr vorkommen? Erkläre anhand eines neuen Schemas.

6. (0.5p) Gegeben sei folgendes Schema. Fremdschlüssel sind kursiv dargestellt, Schlüssel unterstrichen.

Touristen (idTourist, Name, Vorname, *Herkunftsland*)  
 Sehenswürdigkeiten (idSehenswürdigkeit, Name, Beschreibung, *Popularität*)  
 Besucht (idTourist, *idSehenswürdigkeit*, Datum)  
 Touristenroute (idRoute, Länge, Dauer)  
 Gehört (*idSehenswürdigkeit*, *idRoute*)

Formulieren Sie für folgende Anfragen ein Äquivalent in relationaler Algebra:  
 Geben Sie die Sehenswürdigkeiten mit Popularität größer als 8 aus, die zu allen Routen aus der Datenbank gehören.

7. (1p)

a) Erstelle ein B-Baum mit Ordnung 1 und füge der Reihe nach (in der gegebenen Reihenfolge) folgende Werte ein: 20, 51, 70, 85, 90, 75, 72, 23, 30, 22. Zeichne die Struktur des Baumes für jeden Schritt wo sich etwas an der Struktur ändert (man kriegt keine Punkte, wenn man nur die Endstruktur zeigt).  
 b) Lösche den Wert 30 aus dem B-Baum. Zeichne das Endergebnis, und falls nötig auch Zwischenergebnisse. Zeichne jeweils den kompletten Baum.

8. (0.25p) Die folgende Hashdatei benutzt statisches Hashing mit verzahnten Listen. Lösche den Wert 19 aus der Hashdatei.

k	h(k) = k mod 13
16	3
23	10
36	10
25	12
15	2
32	6
29	3
49	10
22	9

2. (0.85p) Gegeben sei folgende Relationen Schema:

Raum (Id, Name)

Vorlesung (Id, Titel, ECTS, ProfessorId)

Professor (Id, Name, Vorname, Geburtsdatum)

Prüfung (VorlesungId, Datum, RaumId)

a) Schreibe eine äquivalente nicht-geschachtelte Abfrage (ohne Unterabfragen) zu der folgenden Abfrage:

```
SELECT T1.nr1 + T2.nr2 FROM
```

```
(SELECT count(*) as nr1 FROM Raum R INNER JOIN Prüfung P ON R.Id = P.RaumId
```

```
INNER JOIN Vorlesung V ON V.Id = P.vorlesungId ) T1,
```

```
(SELECT count(*) as nr2 FROM Vorlesung V2 where V2.Id not in (SELECT VorlesungId  
FROM Prüfung) ) T2
```

b) Nehmen wir an, dass folgende Abfragen oft in der gegebenen Datenbank vorkommen:

```
SELECT * FROM Prüfung WHERE RaumId = ?
```

```
SELECT RaumId, Datum FROM Prüfung WHERE RaumId = ? AND Datum > ?
```

```
SELECT VorlesungId, Datum FROM Prüfung WHERE VorlesungId = ?
```

Schlage zwei Indexe vor, die ihr erstellen würdet um die Anfragen zu optimieren. Für jeden Index bestimme die Art (geclustert oder nicht-geclustert), die Datenstruktur (Baum oder Hash), die Attributen auf welchen der Index erstellt wird, für welche der gegebenen Anfragen der Index nützlich wäre und ob der Index covering ist oder nicht. Begründe die Antwort!

die Art (geclustert oder nicht-geclustert), die Datenstruktur (Baum oder Hash), die Attributen auf welchen der Index erstellt wird, für welche der gegebenen Anfragen der Index nützlich wäre und ob der Index covering ist oder nicht. Begründe die Antwort!

3. (1.75p) Gegeben sei die Relation  $R(A, B, C, D, E, F, G)$  mit folgenden funktionalen Abhängigkeiten:

$B \rightarrow DFG$

$BD \rightarrow CE$

$AC \rightarrow BF$

$DE \rightarrow A$

$E \rightarrow B$

$G \rightarrow EF$

$ADF \rightarrow BC$

a) Gebe die Definition eines Kandidatschlüssels.

b) Finde alle Kandidatschlüssel der Relation R. Begründe.

c) Gebe die Definition der dritten Normalform.

d) Ist die Relation R in 2NF, 3NF und BCNF? Begründe.



Vorname: Thomas Storck

1. (1.05p) Gegeben sei folgende Relation: Lieder (Id, Titel, Artist, Genre, Dauer).  
Gegeben sei folgende Instanz der Relation:

Id	Titel	Artist	Genre	Dauer
1	So What	Ramstein	Pop	2:33
2	Du hast	Metallica	Rock	3:40
3	Whiskey in the Jar	Rammstein	Rock	3:20
4	Zeit	Rammstein	Rock	2:50

Bestimmen Sie für jede der folgenden funktionalen Abhängigkeiten, ob diese für die Daten in der Instanz gilt oder nicht, und begründen Sie Ihre Antwort:

- a) Artist  $\rightarrow$  Genre ✓
- b) Titel, Artist  $\rightarrow$  Genre ✓
- c) Artist  $\rightarrow$  Titel ✗

2. (2.6p) Die Relation KreditZahlungen wird benutzt, um Daten über Kredit Zahlungen zu speichern. Bis zu der Auszahlung eines Kredits macht ein Kunde viele kleinere Zahlungen. Der Kunde wird durch KundeId identifiziert und hat eine eindeutige Adresse. Ein Kunde kann mehrere Kredite haben, aber mit unterschiedlichen Antragsdatums. Ein Kunde kann mehrere Zahlungen am selben Tag haben, aber für unterschiedliche Kredite. Zum Beispiel, ein Kunde hat ein Kredit mit Antragsdatum 1-1-2022 mit dem Betrag 10000 und er macht eine Zahlung mit Summe 2000 am 10.1.2023 (Zahlungsdatum).

- KreditZahlungen (KundeId, KundeName, KundeAdresse, KreditAntragsdatum, Kreditbetrag, Zahlungssumme, Zahlungsdatum)
- a) Gebe ein Kandidatschlüssel der Relation KreditZahlungen.
- b) Ist die Relation KreditZahlungen in BCNF? Begründe die Antwort. Wenn die Relation nicht in BCNF ist, dann zerlege die Relation mithilfe des Zerlegung Korollars (Corollary), sodass die neuen Relationen in BCNF sind.
  - Gebe die Kandidatschlüssel und funktionalen Abhängigkeiten der Relationen an.
  - Begründe warum die Relationen aus der Zerlegung in BCNF sind.
  - Ist die Zerlegung abhängigkeitsbewahrend? Begründe.

3. (1p) Gegeben seien folgende Instanzen der Relationen mit den Schemata (Primärschlüssel sind unterstrichen, die Fremdschlüssel kursiv geschrieben und haben denselben Namen wie die Primärschlüssel, auf die sie sich beziehen):

- Stadt (IdStadt, Name, Land)
  - Wettbewerb (IdWettbewerb, Name, Geburtsdatum, Herkunftsland)
  - Teilnehmer (IdTeilnehmer, Name, Geburtsdatum, Herkunftsland)
  - Preise (IdWettbewerb, IdTeilnehmer, Platz)
- Welches ist das Ergebnis für jede der folgenden Abfragen? Begründen Sie Ihre Antwort!
- a)  $\Pi_{Teilnehmer.Name} (Preise \bowtie_{Preis.IdTeilnehmer=Teilnehmer.IdTeilnehmer} Teilnehmer)$
  - b)  $\Pi_{Preis.IdWettbewerb, Preis.IdTeilnehmer, Preis} \div \Pi_{Wettbewerb.IdWettbewerb} (Wettbewerb \bowtie_{Stadt.Land=Deutschland} Stadt)$

(1.45p) Wir benutzen erweiterbares Hashing für eine Datei, die Datensätze mit folgenden Suchschlüsselwerte enthält:  
 14, 8, 6, 20, 12, 10, 21, 2, 36  
 Zeige die Struktur der Hashdatei (mit lokalen und globalen Trefen) mit Hashfunktion  $h(x) = x \bmod 17$ , wenn man in der  
 gegebenen Reihenfolge die Datensätze einträgt, und die Behälter je 4 Einträge speichern können. Frage an mit lokaler  
 und globaler Tiefe 1.  
 Bitte folge alle Schritte der Reihe nach und zeichne die Struktur für jeden Schritt wo sich diese ändert (man kriegt  
 keine Punkte, wenn man nur die Endstruktur zeigt).

3. (1.5p) Gegeben seien folgende Relationen. Primärschlüssel sind unterstrichen, die Fremdschlüssel kursiv  
 geschrieben und haben denselben Namen wie die Primärschlüssel, auf die sie sich beziehen.  
 Kunden (id, Name, Geburtsdatum, Herkunftsland)  
 Hotels (id, Name, Geburtsdatum, Herkunftsland)  
 Aufenthalts (id, Kunde, Hotel, Anfangsdatum, Enddatum)  
 Schreibe eine Anfrage, welche die durchschnittliche Anzahl der unterschiedlichen Kunden jedes Hotels, mit einem  
 Anfangsdatum des Aufenthalts in dem Zeitraum 2.2.2023 - 3.3.2023, ausgibt. Man darf keine Sichten benutzen.  
 Zum Beispiel, wenn es 3 Hotels gibt, wobei:

- der Hotel H1 hat 1000 Kunden, von denen 200 Kunden mit der gegebenen Bedingung (Anfangsdatum in dem  
 gegebenen Zeitraum)
  - der Hotel H2 hat 500 Kunden, von denen 500 Kunden mit der gegebenen Bedingung
  - und der Hotel H3 hat 500 Kunden, von denen 300 ausgegeben.
6. (1.4p) Für ein Krankenhaus verwaltet man in der gegebenen Tabelle Ärzte. Die Tabelle Ärzte kann keine Null-Werte  
 enthalten. Es kann nicht zwei Ärzte geben mit demselben ArztAusweisid. Mehrere Ärzte können in demselben Büro  
 sitzen. Der untere Snapshot ist nur ein Beispiel, für die folgende Fragen müsst ihr annehmen, dass auch andere Daten  
 eingefügt werden. Für alle Daten gelten jedoch die genannten Einschränkungen.

ArztAusweisid	Name	Vorname	Büro
1452	Tomescu	Andreea	20
1220	Constantin	Mihai	23
3000	Hermann	Anna	34
2300	Popescu	Mihai	20

Begründe für jede der folgenden Aussagen ob die Aussage richtig oder falsch ist. 2. es gibt zwei Ärzte mit dem  
 a) ArztAusweisid ist ein Kandidatschlüssel der Relation (nicht nur für den gegebenen Snapshot)  
 b) (ArztAusweisid, Name, Vorname) ist ein Kandidatschlüssel der Relation (nicht nur für den gegebenen Snapshot)  
 Büro ist ein Kandidatschlüssel der Relation (nicht nur für den gegebenen Snapshot)  
 (Name, Vorname) ist ein Kandidatschlüssel der Relation (nicht nur für den gegebenen Snapshot)

7. Gegeben ist die Relation  $R(A, B, C, D, E)$  mit den funktionalen Abhängigkeiten  $A \rightarrow C, B \rightarrow CD, AC \rightarrow E, D \rightarrow A$ . Welche der folgenden Aussagen ist nicht wahr?

- a)  $A^+ = ACE$
- ☒ b)  $B^+ = ABCD$
- c)  $C^+ = C$
- d)  $D^+ = ACDE$

8. Wann nennen wir eine funktionale Abhängigkeit partiell?

- a) wenn die rechte Seite keine überflüssigen Attribute enthält
- b) wenn die Abhängigkeit nicht trivial ist
- c) wenn die linke Seite kein Superschlüssel ist
- ☒ d) wenn die linke Seite überflüssige Attribute enthält

9. Welche der folgenden Aussagen über geclusterte/gruppierte und nicht-geclusterte/nicht-gruppierte Indexe ist nicht wahr?

- a) es gibt höchstens einen geclusterten Index für eine Datei
- b) ein geclusteter Index ist effizienter für eine Bereichsanfrage als ein nicht-geclusteter Index
- ☒ c) ein Index, der Alternative (2) oder (3) verwendet, ist immer geclustert
- d) ein Index, der Alternative (1) verwendet, ist per Definition geclustert

10. Welche Aussage über den dirty pin ist wahr?

- ☒ a) Der dirty pin markiert ob eine Seite im Pufferplatz geändert wurde oder nicht.
- b) Der dirty pin markiert ob eine Seite im Pufferplatz noch gebraucht wird oder nicht.
- c) Der dirty pin enthält den Verwendungszähler für eine Seite im Pufferplatz.
- d) Der dirty pin markiert ob ein Slot im Pufferplatz leer ist.

II. (0.75p) In dem Disney Park wollen die Veranstalter Statistiken über Besucher und Attraktionen speichern. Gegeben sei das folgende Schema: Attraktionen(id, Name, Beschreibung), Besucher(id, Name, Age), Besucht(besucherId, attraktionId).  
Schreibe eine Anfrage, die die Namen der Attraktionen ausgibt, die von allen Besuchern besucht wurde.

III. (2p) Löse folgende Aufgaben:

a) Definiere die zweite Normalform und Boyce-Codd Normalform. (0.5p)

b) Gebe ein Beispiel von einer Relation, die in der zweiten Normalform ist. Gebe dabei alle Kandidatschlüssel der Relation und die funktionalen Abhängigkeiten an und begründe warum die Relation in der zweiten Normalform ist.

(0.75p)

c) Gebe ein Beispiel von einer Relation (unterschiedlichen Beispiel), die nicht in der zweiten Normalform ist. Gebe dabei alle Kandidatschlüssel der Relation und die funktionalen Abhängigkeiten an und begründe warum die Relation nicht in der zweiten Normalform ist. (0.75p)

2. Gegeben sei das Relationenschema  $\text{Autobesitzer}(\text{besitzerid}, \text{autoid})$ . Welche der folgenden Anfragen gibt die Besitzer aus, die wenigstens zwei Autos besitzen.

- a)  $\pi_{\text{autoid}}(\sigma_{\text{autoid} \neq \text{autoid}}(\text{Autobesitzer}))$   
 b)  $\pi_{\text{autoid}}(\sigma_{\text{autoid} \neq \text{autoid}}(\text{Autobesitzer}))$   
 c)  $\pi_{\text{autoid}}(\sigma_{\text{autoid} \neq \text{autoid}}(\text{Autobesitzer}))$

3. Gegeben ist ein Relationenschema  $R(A, B, C, D, E, F)$  und zwei Mengen von funktionalen Abhängigkeiten:  
 $F_1 = \{AC \rightarrow E, EBF, BAC, CEB, AC \rightarrow D\}$   
 $F_2 = \{AC \rightarrow E, EBF, BAC, CEB, B \rightarrow DF\}$   
 Sind  $F_1$  und  $F_2$  äquivalent?

- a) Man kann nicht wissen ohne die Semantik der Attribute zu kennen.  
 b) Ja  
 c) Nein

4. Für einen Projekt werden Studenten in Teams eingeteilt und jedes Team kriegt ein Thema. Man nimmt an, dass ein Student zu einem einzigen Team gehört. Wie muss die Beziehung zwischen den Relationen *Studenten* und *Teams* modelliert werden?
- a) ein Fremdschlüssel in der Relation *Studenten*, der auf *Teams* verweist  
 b) ein Fremdschlüssel in der Relation *Teams*, der auf *Studenten* verweist  
 c) eine neue Relation  
 d) ein Fremdschlüssel in der Relation *Studenten*, der auf *Teams* verweist, und ein Fremdschlüssel in der Relation *Teams*, der auf *Studenten* verweist

5. Worauf bezieht sich die logische Datenunabhängigkeit?
- a) einfache Änderungen des Datenbankschemas sollen keine Auswirkungen auf die Zugriffstechniken und auf die Art der Datenspeicherung haben  
 b) Änderungen an der Art der Datenspeicherung und der Zugriffstechniken sollen keine Auswirkungen auf Anwendungsprogramme haben  
 c) Änderungen in der Anwendung (Application) sollen keine Auswirkungen auf das Datenbankschema haben  
 d) einfache Änderungen des Datenbankschemas sollen keine Auswirkungen auf die Anwendung haben

6. Gegeben sei die folgende Ausprägung der Relation  $R(A, B, C)$ :

A	B	C
Ana	F1	9
Ioana	NULL	7
Andrei	F2	6
Mihai	F1	6

Welche der folgenden Anfragen ist korrekt und gibt das maximale Wert aus:

- a)  $\text{select sum(C)/count(B) - count(distinct C) from R}$   
 b)  $\text{select sum(C)/avg(C) from R}$   
 c)  $\text{select sum(C)/count(distinct C) from R}$   
 d)  $\text{select avg(C) - count(distinct C) from R}$



C

# Prüfung Datenbanken

$\frac{I}{II} 1,60$   
 $\frac{II}{IV} 0$   
 $\frac{IV}{V} 1,50$   
 $\frac{V}{VI} 0,50$   
 $\frac{VI}{VII} 0,65$   
 $\frac{VII}{VIII} 1,25$   
 $\frac{VIII}{IX} 1,00$   
 $\frac{IX}{X} 5,30$

Name: Pop Andreea  
 Gruppe: 722  
 Datum: 24.01.2019

1. (2p = 10 x 0.2p) Folgende Fragen haben eine einzige richtige Antwort. Trage alle Antworten in der folgenden Tabelle ein!!

1	b	6	c
2	b	7	b
3	c	8	d
4	a	9	c
5	b	10	a

1. Die folgende Hashdatei benutzt statisches Hashing mit verzahnten Listen.

k	h(k) = k mod 13
16	3
23	10
36	10
25	12
19	6
32	6
29	3
49	10
32	5
23	12
25	7
36	11

Nach dem Löschen des Wertes 16 sieht die Hashdatei folgendermaßen aus:

a) b) c) d)

0		0		0		0	
1		1		1		1	
2		2		2		2	
3		3	16	3	29	3	29
4		4		4		4	
5	22	5	22	5	22	5	22
6	19	6	19	6	19	6	19
7	49	7	49	7	49	7	49
8	29	8	29	8	29	8	29
9	32	9	32	9	32	9	32
10	23	10	23	10	23	10	23
11	25	11	25	11	25	11	25
12	36	12	36	12	36	12	36



ii

Attraktionen (id, Name, ~~Name~~, Beschreibung)  
Besucher (id, Name, Age)  
Besucht (besucherid, attraktionid)

SELECT A. Name  
FROM Attraktionen A  
WHERE A.id = ALL (SELECT B.attraktionid  
FROM Besucht B)

2. ADB - cheie primară

Pop Andrei

3. Die Relation ist nicht in BCNF, weil ADB nicht eine Superschlüssel ist.

(A, B, C, D, E)  $B \rightarrow E$  (B, E), (A, B, C, D)  
 $E \rightarrow C$  (B, E, C), (A, B, D, E) ?  
 $D \rightarrow C$  (D, C), (A, B, D, E)  
 $E \rightarrow B$  (E, B), (A, C, D, E)

V1 Besucht (Kneipe, Gast)

Kneipe	Gast
Klausen	Haus
Klausen	Ana
Insomnia	Haus

Trinkt (Gast, Bier)

Gast	Bier
Haus	Pils
Ana	Hefeweizen
Haus	Hefeweizen

0,25

b) verlustlos - atunci când după înlocuirea unui Tupel în  
 este posibil să se adăgească.

abhängigkeitsbewahrend - când dependențele funcționale se păstrează

c) Nein, pt că în Tupel se pierde (pierde)

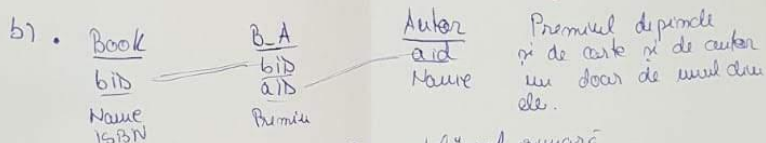
Insomnia, Haus  
 Insomnia, Haus

și Haus, Pils  
 Haus, Pils

?  
 0

- III. a). 2 Normalform <sup>sa chere</sup>  
 - toate atributele trebuie să depindă de mbraga chere, adică să nu existe dependențe parțiale.

- 0,5 BC Normalform  
 - toate dependențele sunt date de chere  
 -  $\alpha \rightarrow A$ , ~~cu~~  $\alpha$  este un Superschlüssel, oder  $\alpha \rightarrow A$  trivial



•  $R(A, B, C, D)$  AB - schlüssel primar

- 0,25  $AB \rightarrow C$   $C \rightarrow AB$  ~~triviale~~  
 $AB \rightarrow D$   $D \rightarrow AB$

- este în 2NF pt că atributele C și D depind de mbraga chere AB nu doar de A sau B.

- Kandidatschlüssel:  $(AB)^+ = ABCD$

- c).  $R(A, B, C, D, E)$  AB - schlüssel primar 0,75  
 $C \rightarrow B$  }  $\left. \begin{array}{l} \text{dependente parțiale} \\ \text{dependente totale} \end{array} \right\} \Rightarrow$  nu în 2NF  
 $E \rightarrow A$   
 $CD \rightarrow AB$

Kandidatschlüssel:  $(CD)^+ = CDBAE$

~~Person~~  
~~Car~~  
~~PersonID~~  
~~Preis~~  
~~Stärke~~  
~~Stärke~~

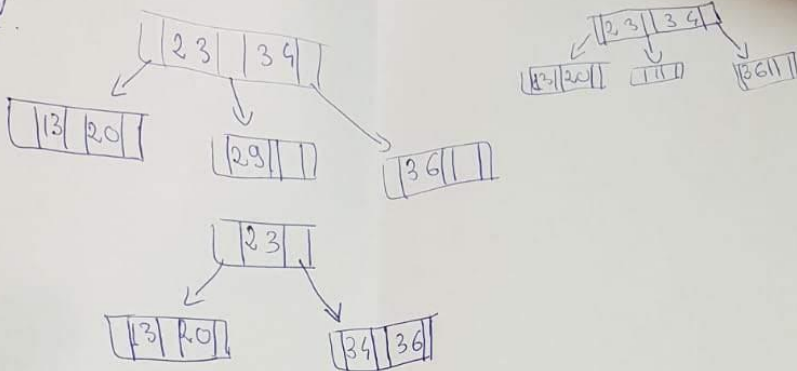
Car  
 CID  
 Name  
 Color

Car-Person  
 CarID  
 PersonID  
 Preis

Person  
 PID  
 Name  
 Age

Prefatul depinde doar de consiliu, nu depinde ni de persoane  
 $\Rightarrow$  'dependență parțială'  $\Rightarrow$  nu e în 2NF

IV.



$R(A, B, C, D, E)$   $F = \{B \rightarrow E, BE \rightarrow C, D \rightarrow C, E \rightarrow B\}$

1.  $B \rightarrow E$   
 $BE \rightarrow C$   
 $D \rightarrow C$   
 $E \rightarrow B$

$A^+ = A \Rightarrow$  nu se poate afla de nimic de aici  
trebuie să oporționez ceva

$(AB)^+ = ABE$  X  
 $D^+ = DC$  - ni D trebuie să oporționez ceva

$(ADB)^+ = ABCDE$  ✓  $AD^+ = ?$   
0,45

$(AE)^+ = AEB$  X  
 $(ADE)^+ = ABCDE$  ✓  
 $(ADC)^+ = ADC$  X

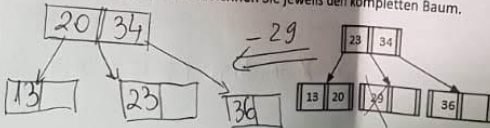
$B)^+ = ADCBE$  ✓  
 $CE)^+ = ADCEB$  ✓

Candidatatschlüssel:  $ADB, ADE, \underline{ADCB}, \underline{ADCE}$  ??



- IV. (0.75p) In dem Disney Park wollen die Veranstalter Statistiken über Besucher und Attraktionen speichern. Gegeben sei das folgende Schema: Attraktionen(id, Name, Beschreibung), Besucher(id, Name, Age), Besucht(besucherId, attraktionId). Schreibe eine Anfrage, die die Namen der Attraktionen ausgibt, die von allen Besuchern besucht wurde.

- V. (0.5p) Lösche den Wert 29 aus den gegebenen B-Baum mit Ordnung 1. Zeichnen Sie das Endergebnis, und falls nötig auch Zwischenergebnisse. Zeichnen Sie jeweils den kompletten Baum.



- VI. (2.25p) Gegeben sei das Relationenschema  $R(A, B, C, D, E)$  und die Menge der funktionalen Abhängigkeiten  $F = \{B \rightarrow E, BE \rightarrow C, D \rightarrow C, E \rightarrow B\}$

- Bestimme alle Kandidatschlüssel der Relation R. Erkläre dabei warum diese Kandidatschlüssel sind und warum es keine andere gibt. (0.5p)
- Beweise/Erkläre ob die Relation in der zweiten und in der dritten Normalform ist. (0.75p)
- Beweise/Erkläre ob die Relation in BCNF ist oder nicht. Falls die Relation nicht in BCNF ist, zerlege diese (verlustlos) in mehreren BCNF Relationen. Erkläre wie ihr zu der Zerlegung gekommen seid. (1p)

1. Ich berechne die Hülle von  $A^+$ , weil  $A \in FDs \Rightarrow$

$$\rightarrow A^+ = A^\#$$

$$A^+ = ABEC$$

$$AB^+ = ABEC$$

$$AC^+ = AC$$

$$AD^+ = ADC$$

$$AE^+ = AEB$$

$$DB^+ = DBEC$$

geg berechne ich  
e von D:  $D^+ = DC$

$$AD^+ = ADC$$

$$ADB^+ = ADBEC \text{ (KS)}$$

$$ADE^+ = ADE$$

$$ADE^+ = ADEBC \text{ (KS)}$$

$$KS: ADB, ADE$$

$$\begin{aligned} 2. B &\rightarrow E \\ BE &\rightarrow C \\ D &\rightarrow C \\ E &\rightarrow B \end{aligned}$$

- Die Formel ist nicht im 2NF wegen  $D \rightarrow C$ .
- Die Formel ist nicht im 3NF wegen  $D \rightarrow C$ , weil C ist nicht prim.

Wird sowohl A, als auch D sich in der Hülle befinden sollen suche und berechne ich nur die, die minimal sind und A, D enthalten.

Ab hier sind die KS nicht mehr minimal

warum?

Worauf bezieht sich die logische Datenunabhängigkeit?

- a) einfache Änderungen des Datenbankschemas sollen keine Auswirkungen auf die Zugriffstechniken und auf die Art der Datenspeicherung haben
- b) Änderungen an der Art der Datenspeicherung und der Zugriffstechniken sollen keine Auswirkungen auf Anwendungsprogramme haben
- c) einfache Änderungen des Datenbankschemas sollen keine Auswirkungen auf die Anwendung haben
- d) Änderungen in der Anwendung (application) sollen keine Auswirkungen auf das Datenbankschema haben

9. Wann nennen wir eine funktionale Abhängigkeit partiell?  $\rightarrow x \rightarrow y$  no dan  $\exists x \subset x \rightarrow y$

- a) wenn die linke Seite überflüssige Attribute enthält
- b) wenn die Abhängigkeit nicht trivial ist
- c) wenn die linke Seite kein Superschlüssel ist
- d) wenn die rechte Seite keine überflüssigen Attribute enthält

10. Für einen Projekt werden Studenten in Teams eingeteilt und jedes Team kriegt ein Thema. Man nimmt an, dass ein Student zu einem einzigen Team gehört. Wie muss die Beziehung zwischen den Relationen Studenten und Teams modelliert werden?

- a) ein Fremdschlüssel in der Relation Teams, der auf Studenten verweist
- b) eine neue Relation
- c) ein Fremdschlüssel in der Relation Studenten, der auf Teams verweist
- d) ein Fremdschlüssel in der Relation Studenten, der auf Teams verweist, und ein Fremdschlüssel in der Relation Teams, der auf Studenten verweist

11. (1.5p) Nach einer Kneipentour speichert man folgende Informationen über Biertrinker in der Tabelle Biertrinker(Kneipe, Gast, Bier)

Kneipe	Gast	Bier
Klausen	Hans	Pils
Klausen	Anna	Hefeweizen
Insomnia	Hans	Hefeweizen
Insomnia	Hans	Pils

Man entscheidet, die Relation Biertrinker folgendermaßen zu zerlegen: Besucht(Kneipe, Gast) und Trinkt(Gast, Bier). Löse folgende Aufgaben:

- a) Welche Tupel gehören zu den Tabellen aus der Zerlegung? Zeige die Ausprägungen. (0.25p)
- b) Wann heißt eine Zerlegung verlustlos? Wann heißt eine Zerlegung abhängigkeitsbewahrend? (0.5p)
- c) Ist die gegebene Zerlegung verlustlos? Begründe. (0.75p) *Nu*

12. (2p) Löse folgende Aufgaben:

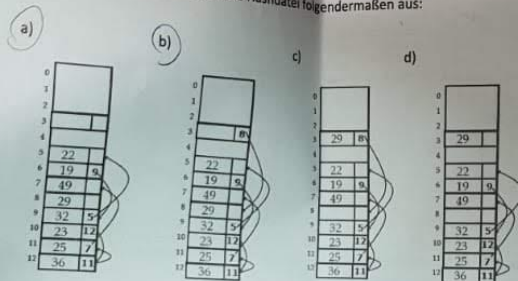
- a) Definiere die zweite Normalform und Boyce-Codd Normalform. (0.5p)
- b) Gebe ein Beispiel von einer Relation, die in der zweiten Normalform ist. Gebe dabei alle Kandidatschlüssel der Relation und die funktionalen Abhängigkeiten an und begründe warum die Relation in der zweiten Normalform ist. (0.75p)
- c) Gebe ein Beispiel von einer Relation (unterschiedlichen Beispiel), die nicht in der zweiten Normalform ist. Gebe dabei alle Kandidatschlüssel der Relation und die funktionalen Abhängigkeiten an und begründe warum die Relation nicht in der zweiten Normalform ist. (0.75p)

Fach ID FachName

4. Die folgende Hashdatei benutzt statisches Hashing mit verzahnten Listen.

$k$	$h(k) = k \bmod 13$
16	3
23	10
36	10
25	12
19	6
32	6
29	3
49	10
22	9

Nach dem Löschen des Wertes 16 sieht die Hashdatei folgendermaßen aus:



Welche Aussage über den dirty pin ist wahr?

- a) Der dirty pin markiert ob eine Seite im Pufferplatz noch gebraucht wird oder nicht.  
 b) Der dirty pin markiert ob eine Seite im Pufferplatz geändert wurde oder nicht.  
 c) Der dirty pin enthält den Verwendungszähler für eine Seite im Pufferplatz.  
 d) Der dirty pin markiert ob ein Slot im Pufferplatz leer ist.

6. Gegeben ist die Relation  $R(A, B, C, D, E)$  mit den funktionalen Abhängigkeiten  $A \rightarrow C$ ,  $B \rightarrow CD$ ,  $AC \rightarrow E$ ,  $D \rightarrow A$ . Welche der folgenden Aussagen ist nicht wahr?

- a)  $B^+ = ABCD$   
 b)  $A^+ = ACE$   
 c)  $C^+ = C$   
 d)  $D^+ = ACDE$

$$C^+ = BCDAE$$

$$D^+ = DACAE$$

7. Welche der folgenden Aussagen über geclusterte/gruppierte und nicht-geclusterte/nicht-gruppierte Indexte ist nicht wahr?

- a) es gibt höchstens einen geclusterten Index für eine Datei  
 b) ein geclusterter Index ist effizienter für eine Bereichsanfrage als ein nicht-geclusterter Index  
 c) ein Index, der Alternative (1) verwendet, ist per Definition geclustert  
 d) ein Index, der Alternative (2) oder (3) verwendet, ist immer geclustert

# Prüfung Datenbanken

Name: Piedra Anamaria

Gruppe: 723

Datum: 24.01.2019

2p = 10 x 0.2p) Folgende Fragen haben eine einzige richtige Antwort. Trage alle Antworten in der folgenden Tabelle ein!!

1	<u>b</u>	6	<u>a</u>	✓
2	<u>b</u>	7	<u>c</u>	
3	<u>a</u>	8	<u>c</u>	✓
4	<u>b</u>	9	<u>c</u>	
5	<u>b</u>	10	<u>c</u>	✓

Gegeben ist ein Relationenschema  $R(A,B,C,D,E,F)$  und zwei Mengen von funktionalen Abhängigkeiten:  
 $F_1 = \{AC \rightarrow AE, EBF, BAC, CEB, AC \rightarrow D\}$   
 $F_2 = \{AC \rightarrow AE, EBF, BAC, CEB, B \rightarrow DF\}$

Sind  $F_1$  und  $F_2$  äquivalent?

- a) Ja
- b) Nein
- c) Man kann nicht wissen ohne die Semantik der Attribute zu kennen.

2. Gegeben sei das Relationenschema Autobesitzer(besitzerId, autoid). Welche der folgenden Anfragen gibt die Besitzer aus, die wenigstens zwei Autos besitzen.

- a)  $\rho_{A_1}(\text{Autobesitzer}) \bowtie_{A_1.\text{autoid} < A_2.\text{autoid}} (\pi_{A_1.\text{besitzerId}}(\sigma_{A_1.\text{autoid} < A_2.\text{autoid}}(A_1 \bowtie_{A_1.\text{besitzerId} = A_2.\text{autoid}} A_2)))$
- b)  $\rho_{A_1}(\text{Autobesitzer}) \bowtie_{A_1.\text{autoid} < A_2.\text{autoid}} (\pi_{A_1.\text{besitzerId}}(\sigma_{A_1.\text{autoid} < A_2.\text{autoid}}(A_1 \bowtie_{A_1.\text{besitzerId} = A_2.\text{autoid}} A_2)))$
- c)  $\rho_{A_1}(\text{Autobesitzer}) \bowtie_{A_1.\text{autoid} < A_2.\text{autoid}} (\pi_{A_1.\text{besitzerId}}(\sigma_{A_1.\text{autoid} < A_2.\text{autoid}}(A_1 \bowtie_{A_1.\text{besitzerId} = A_2.\text{autoid}} A_2)))$

3. Gegeben sei die folgende Ausprägung der Relation  $R(A,B,C)$ :

A	B	C
Ana	F1	9
Ioana	NULL	7
Andrei	F2	6
Mihai	F1	6

Welche der folgenden Anfragen ist korrekt und gibt das maximale Wert aus:

- a)  $\text{select sum(C)/count(distinct C) from R}$
- b)  $\text{select sum(C)/avg(C) from R}$
- c)  $\text{select sum(C)/(count(B) - count(distinct C)) from R}$
- d)  $\text{select avg(C) - count(distinct C) from R}$

$$\begin{aligned}
 &S = 12 + 12 = 13 + 9 = 28 \\
 &28/3 = 9 \frac{2}{3} \\
 &28/4 = 7 \\
 &\frac{28}{6} \times 4 = 18 \frac{2}{3}
 \end{aligned}$$



III. (2.25p) Gegeben sei das Relationenschema  $R(A, B, C, D, E)$  und die Menge der funktionalen Abhängigkeiten

$F = \{D \rightarrow B, BE \rightarrow C, A \rightarrow C, B \rightarrow D\}$

- Bestimme alle Kandidatschlüssel der Relation  $R$ . Erkläre dabei warum diese Kandidatschlüssel sind und warum es keine andere gibt. (0.5p)
- Beweise/Erkläre ob die Relation in der zweiten und in der dritten Normalform ist. (0.75p)
- Beweise/Erkläre ob die Relation in BCNF ist oder nicht. Falls die Relation nicht in BCNF ist, zerlege diese (verlustlos) in mehreren BCNF Relationen. Erkläre wie ihr zu der Zerlegung gekommen seid. (1p)

IV. (1.5p) Die Cinemas aus den Malls speichern Informationen über Kunden und die Filme in der Tabelle  
Cinemas(Cinema, Kunde, Film)

Cinema	Kunde	Film
Cinema City	Anna	Papillon
Polus	Hans	Papillon
Cinema City	Jürgen	Glas
Cinema City	Anna	Aquaman

Man entscheidet, die Tabelle folgendermaßen zu zerlegen:  $Spielt(Cinema, Film)$  und  $Kunden(Cinema, Kunde)$ . Löse folgende Aufgaben:

- Welche Tupel gehören zu den Tabellen aus der Zerlegung? Zeige die Ausprägungen. (0.25p)
- Wann heißt eine Zerlegung verlustlos? Wann heißt eine Zerlegung abhängigkeitsbewahrend? (0.5p)
- Ist die gegebene Zerlegung verlustlos? Begründe. (0.75p)

V. (2p) 2. Löse folgende Aufgaben:

- Definiere die erste Normalform und die dritte Normalform. (0.5p)
- Gebe ein Beispiel von einer Relation, die in der dritten Normalform ist. Gebe dabei alle Kandidatschlüssel der Relation und die funktionalen Abhängigkeiten an und begründe warum die Relation in der dritten Normalform ist. (0.75p)
- Gebe ein Beispiel von einer Relation (unterschiedlichen Beispiel), die nicht in der dritten Normalform ist. Gebe dabei alle Kandidatschlüssel der Relation und die funktionalen Abhängigkeiten an und begründe warum die Relation nicht in der dritten Normalform ist. (0.75p)

VI. (0.75p) Gegeben sei das folgende Schema, die Informationen speichert über Kunden und Filme, die sie ausgeliehen haben: Kunde(id, Name), Filme(id, Titel), Leiht(kundeId, filmId)  
Schreibe eine Anfrage, die die Namen der Kunden ausgibt, die alle Filme ausgeliehen haben.

Die folgende Hashdatei benutzt statisches Hashing mit verzahnten Listen.

0		
1		
2	15	12
3		
4	17	11
5		
6		
7		
8		
9	14	
10	23	
11	30	
12	28	9

k	$h(k) = k \bmod 13$
17	4
23	10
15	2
28	2
30	4

Nach dem Einfügen des Wertes 41 sieht die Hashdatei folgendermaßen aus:

a)

0		
1		
2	41	12
3		
4	17	11
5		
6		
7		
8		
9		
10	23	
11	30	
12	28	9

b)

0	41	
1		
2	15	12
3		
4	17	11
5		
6		
7		
8		
9		
10	23	
11	30	
12	28	9

c)

0		
1		
2	41	12
3		
4	17	11
5		
6		
7		
8		
9	28	
10	23	
11	30	
12	15	9

d)

0		
1		
2	15	12
3		
4	17	11
5		
6		
7		
8		
9	41	
10	23	
11	30	
12	28	9

II. (0.5p) Lösche den Wert 26 aus den gegebenen B-Baum mit Ordnung 1. Zeichnen Sie das Endergebnis, und falls nötig auch Zwischenergebnisse. Zeichnen Sie jeweils den kompletten Baum.



5. Gegeben sei die folgende Ausprägung der Relation  $R(A,B,C)$ :

A	B	C
Ana		9
Ioana	F1	7
Andrei	NULL	6
Mihai	F2	6
	F1	

Welche der folgenden Anfragen ist korrekt und gibt das maximale Wert aus:

- a) `select avg(C) - count(B) from R`  $7 - 3 = 4$   
b) `select avg(count(C)) from R`  $4$   
c) `select avg(C) - count(distinct B) from R`  $7 - 3 = 4$   
d) `select max(C) - avg(C) from R`  $9 - 7 = 2$
6. Gegeben sei das Relationenschema  $R(A,B,C,D,E)$  mit der Menge der funktionalen Abhängigkeiten  $F = \{A \rightarrow C, B \rightarrow AC, C \rightarrow E\}$ . Welche der folgenden Aussagen ist wahr?  
a)  $A^+ = ABCE$   
b)  $A^+ = ACDE$   
c)  $A^+ = ACE$   
d)  $A^+ = ABCDE$
7. Welche ist keine mögliche Ursache für Anomalien in einer Relation:  
a) falsche Fremdschlüssel  
b) Redundanz innerhalb der Relation  
c) Vermischung zweier Entitätstypen
8. Wann ist das Attribut einer Relation prim?  
a) wenn das Attribut nur in der linken Seite der funktionalen Abhängigkeiten vorkommt  
b) wenn das Attribut nur in der rechten Seite der funktionalen Abhängigkeiten vorkommt  
c) wenn das Attribut eindeutig ist  
d) wenn das Attribut in einem Kandidatschlüssel enthalten ist
9. Eine Firma will die Angestellte und die Projekte, an denen diese arbeiten, speichern. Man nimmt an, dass ein Angestellter an mehreren Projekten gleichzeitig arbeiten kann. Wie muss die Beziehung zwischen den Relationen Angestellte und Projekte modelliert werden?  
a) ein Fremdschlüssel in der Relation Angestellte, der auf Projekte verweist  
b) eine neue Relation  
c) ein Fremdschlüssel in der Relation Projekte, der auf Angestellte verweist  
d) ein Fremdschlüssel in der Relation Angestellte, der auf Projekte verweist, und ein Fremdschlüssel in der Relation Projekte, der auf Angestellte verweist

# Prüfung Datenbanken

Name: Faridat Sonja  
 Gruppe: #22  
 Datum: 24.01.2019

(2p = 10 x 0.2p) Folgende Fragen haben eine einzige richtige Antwort. Trage alle Antworten in der folgenden Tabelle ein!!

1	<u>C</u>	6	<u>C</u>
2	<u>B</u>	7	<u>A</u>
3	<u>A</u>	8	<u>D</u>
4	<u>D</u>	9	<u>B</u>
5	<u>C</u>	10	<u>D</u>

1. Gegeben ist ein Relationenschema  $R(A,B,C,D,E,F)$  und zwei Mengen von funktionalen Abhängigkeiten:

$F_1 = \{ABE, ECD, CD AC, ADE, CD F\}$

$F_2 = \{ABE, ECD, CD AC, ADE, E BF\}$

Sind  $F_1$  und  $F_2$  äquivalent?

- a) Man kann nicht wissen ohne die Semantik der Attribute zu kennen.
- b) Ja
- c) ☒ Nein

2. Ein Autohersteller bietet dasselbe Auto in unterschiedlichen Farben an. Diese Informationen werden in dem Relationenschema  $Autofarbe(autoid, farbe)$  gespeichert. Welche der folgenden Anfragen gibt die Autos an, welche in roter und in grüner Farbe angeboten werden.

- a)  $\pi_1(Autofarbe), \sigma_{farbe = 'rot' \vee farbe = 'grün'}(F)$
- b)  $\pi_1(Autofarbe), \sigma_{farbe = 'rot' \vee farbe = 'grün'}(F)$
- c)  $\pi_1(Autofarbe), \sigma_{farbe = 'rot' \vee farbe = 'grün'}(F)$

3. Welche der folgenden Aussagen beschreibt die Zugriffsbewegungszeit (seek time) am besten?

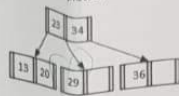
- a) Die Zeit um den Schreib/LeseKopf auf der richtigen Spur zu positionieren
- b) Die Zeit bis den gesuchten Block am Kopf vorbeigeführt wird
- c) Die Zeit um die Daten von der Plattenoberfläche zu lesen oder auf die Plattenoberfläche zu schreiben
- d) ☒ Die durchschnittliche Zeit um eine bestimmte Seite zu finden

4. Welche der folgenden Aussagen ist keine gute Regel in der Praxis?

- a) Erstelle nicht-geclusterte Indexe für Fremdschlüsseln
- b) Jede Tabelle sollte einen geclusterten Index haben.
- c) Bei Indexen mit zusammengesetzten Schlüsseln, ordne die Spalten so, dass die am meisten selektiven die ersten sind
- d) ☒ Erstelle nicht-geclusterte Indexe für alle Attribute, die in einem WHERE-Klausel vorkommen.



- IV. (0.5p) Lösche den Wert 29 aus den gegebenen B-Baum mit Ordnung 1. Zeichnen Sie das Endergebnis, und falls nötig auch Zwischenergebnisse. Zeichnen Sie jeweils den kompletten Baum.



- V. (2.25p) Gegeben sei das Relationenschema  $R(A, B, C, D, E)$  und die Menge der funktionalen Abhängigkeiten  $F = \{B \rightarrow E, BE \rightarrow C, D \rightarrow C, E \rightarrow B\}$
1. Bestimme alle Kandidatschlüssel der Relation R. Erkläre dabei warum diese Kandidatschlüssel sind und warum es keine andere gibt. (0.5p)
  2. Beweise/Erkläre ob die Relation in der zweiten und in der dritten Normalform ist. (0.75p)
  3. Beweise/Erkläre ob die Relation in BCNF ist oder nicht. Falls die Relation nicht in BCNF ist, zerlege diese (verlustlos) in mehreren BCNF Relationen. Erkläre wie ihr zu der Zerlegung gekommen seid. (1p)

- VI. (1.5p) Nach einer Kneipentour speichert man folgende Informationen über Biertrinker in der Tabelle Biertrinker(Kneipe, Gast, Bier)

Kneipe	Gast	Bier
Klausen	Hans	Pils
Klausen	Anna	Hefeweizen
Insomnia	Hans	Hefeweizen
Insomnia	Hans	Pils

Man entscheidet, die Relation Biertrinker folgendermaßen zu zerlegen:  $Besucht(Kneipe, Gast)$  und  $Trinkt(Gast, Bier)$ . Löse folgende Aufgaben:

- a) Welche Tupel gehören zu den Tabellen aus der Zerlegung? Zeige die Ausprägungen. (0.25p)
- b) Wann heißt eine Zerlegung verlustlos? Wann heißt eine Zerlegung abhängigkeitsbewahrend? (0.5p)
- c) Ist die gegebene Zerlegung verlustlos? Begründe. (0.75p)

- 7) Gegeben sei eine Relation  $R(A, B, C, D, E)$  mit folgenden funktionalen Abhängigkeiten:
- $A \rightarrow C, AC \rightarrow D, AB \rightarrow E, BCD \rightarrow AE$
- Welche der folgenden Aussagen sind wahr?
- a)  $AC \rightarrow D$  ist eine partielle Abhängigkeit.
  - b)  $BCD \rightarrow CD$  ist eine triviale Abhängigkeit.
  - c)  $BCD \rightarrow AE$  ist eine volle Abhängigkeit.
  - d)  $AB$  ist ein Kandidatschlüssel.
  - e)  $AC$  ist ein Superschlüssel.

Für folgende Aufgaben gebe vollständige Lösungen und erkläre die Schritte, die Ihr in den Lösungen folgt.

1. (0.5p)
  - a) Definiere die Referenz-Integritätsregel und erläutere die Definition anhand eines Beispiels.
  - b) Definiere eine funktionale Abhängigkeit und erläutere die Definition anhand eines Beispiels.
2. (0.85p) Gegeben sei folgende Relationen Schema:
  - Raum (Id, Name)
  - Vorlesung (Id, Titel, ECTS, Professorid)
  - Professor (Id, Name, Vorname, Geburtsdatum)
  - Prüfung (Vorlesungid, Datum, Raumid)
  - a) Schreibe eine äquivalente nicht-geschachtelte Abfrage (ohne Unterabfragen) zu der folgenden Abfrage:
 

```
SELECT T1.nr1 + T2.nr2 FROM
(SELECT count(*) as nr1 FROM Raum R INNER JOIN Prüfung P ON R.Id = P.Raumid
INNER JOIN Vorlesung V ON V.Id = P.vorlesungid ) T1,
(SELECT count(*) as nr2 FROM Vorlesung V2 where V2.Id not in (SELECT Vorlesungid
FROM Prüfung ) T2
```
  - b) Nehmen wir an, dass folgende Abfragen oft in der gegebenen Datenbank vorkommen:
 

```
SELECT * FROM Prüfung WHERE Raumid = ?
SELECT Raumid, Datum FROM Prüfung WHERE Raumid = ? AND Datum > ?
SELECT Vorlesungid, Datum FROM Prüfung WHERE Vorlesungid = ?
```

Schlage zwei Indexe vor, die ihr erstellen würdet um die Anfragen zu optimieren. Für jeden Index bestimme die Art (geclustert oder nicht-geclustert), die Datenstruktur (Baum oder Hash), die Attributen auf welchen der Index erstellt wird, für welche der gegebenen Anfragen der Index nützlich wäre und ob der Index covering ist oder nicht. Begründe die Antwort!

3. (1.75p) Gegeben sei die Relation  $R(A, B, C, D, E, F, G)$  mit folgenden funktionalen Abhängigkeiten:
  - $B \rightarrow DFG$
  - $BD \rightarrow CE$
  - $AC \rightarrow BF$
  - $DE \rightarrow A$
  - $E \rightarrow B$
  - $G \rightarrow EF$
  - $ADF \rightarrow BC$
  - a) Gebe die Definition eines Kandidatschlüssels.
  - b) Finde alle Kandidatschlüssel der Relation  $R$ . Begründe.
  - c) Gebe die Definition der dritten Normalform.
  - d) Ist die Relation  $R$  in 2NF, 3NF und BCNF? Begründe.

3. a) Kandidat  
attribut  
este G  
de min  
au alea

b)  $\{B\} = \{A, B\}$   
(pt. co)  $B \rightarrow$

$\{A\} = \{A, B\}$   
(pt. co)  $E \rightarrow$

$\{G\} = \{A, B\}$   
(pt. co)  $G \rightarrow$

$\Rightarrow$  Kandidat  
Sum Sum

c) Eine Relat  
FDs F is  
für alle Ab

1. B
2. A
3. B

4. (1.75p) Gegeben sei die Relation  $R(A, B, C, D, E, F, G)$  mit folgenden funktionalen Abhängigkeiten:  
 $F = \{ABG \rightarrow E, DG \rightarrow B, B \rightarrow C, G \rightarrow D, C \rightarrow BD, AG \rightarrow BC, B \rightarrow D\}$   
 Ist die Relation  $R$  in der dritten Normalform? Begründe. Wenn die Relation nicht in 3NF ist, dann benutze das Syntheselgorithmus um eine 3NF Zerlegung zu finden. Erkläre die Schritte, die ihr folgt. Ist diese Zerlegung verlustlos und abhängigkeitsbewahrend? Begründe.

5. (1p) Gebe ein Beispiel aus der Realität (das auch Sinn macht), unterschiedlich von den Beispielen aus der Vorlesung und Seminar, von einer Relation, die in der zweiten Normalform, aber nicht in der dritten Normalform ist. Erkläre warum diese Relation nicht in der dritten Normalform ist. Erkläre was für Probleme in eurem Schema vorkommen können und gebe ein Beispiel (mit einem Snapshot der Tabelle) wann die beschriebenen Probleme vorkommen. Wie würdet ihr die Tabelle modellieren, sodass die beschriebenen Probleme nicht mehr vorkommen? Erkläre anhand eines neuen Schemas.

6. (0.5p) Gegeben sei folgendes Schema. Fremdschlüssel sind kursiv dargestellt, Schlüssel unterstrichen.

Touristen (idTourist, Name, Vorname, Herkunftsland)  
 Sehenswürdigkeiten (idSehenswürdigkeit, Name, Beschreibung, Popularität)  
 Besucht (idTourist, idSehenswürdigkeit, Datum)  
 Touristenroute (idRoute, Länge, Dauer)  
 Gehört (idSehenswürdigkeit, idRoute)

Formulieren Sie für folgende Anfragen ein Äquivalent in relationaler Algebra:  
 Geben Sie die Sehenswürdigkeiten mit Popularität größer als 8 aus, die zu allen Routen aus der Datenbank gehören.

7. (1p)  
 a) Erstelle ein B-Baum mit Ordnung 1 und füge der Reihe nach (in der gegebenen Reihenfolge) folgende Werte ein: 20, 51, 70, 85, 90, 75, 72, 23, 30, 22. Zeichne die Struktur des Baumes für jeden Schritt wo sich etwas an der Struktur ändert (man kriegt keine Punkte, wenn man nur die Endstruktur zeigt).  
 b) Lösche den Wert 30 aus dem B-Baum. Zeichne das Endergebnis, und falls nötig auch Zwischenergebnisse. Zeichne jeweils den kompletten Baum.

8. (0.25p) Die folgende Hashdatei benutzt statisches Hashing mit verzahnten Listen. Lösche den Wert 19 aus der Hashdatei.

3	16	8
4	23	
5	19	6
6	10	
7	20	
8	30	25
9	25	7
10	50	11

k	h(k) = k mod 11
16	5
23	10
50	6
25	3
10	10
6	6
30	9
25	3
7	6
50	6

## Prüfung Datenbanken

LUNG

3. a) Kandidatschlüssel - O das kandidat a attribute einer Relation R. in case für este e superschlüssel) ? theore so de minimalitate (adica numai subse aleasa nu poate fi la kandidat lui

b)  $\{B\}_+ = \{A, B, C, D, E, F, G\}$  - es e cheia cano a relatiei R  
 (pt. cō  $B \rightarrow AFG, G \rightarrow EF, DE \rightarrow A, B \rightarrow C$

$\{E\}_+ = \{A, B, C, D, E, F, G\}$  - cheia kandidat p! pentru R(A,B,C,D,E,F,G)  
 (pt. cō  $E \rightarrow B$  și am arătat cō B e superschlüssel

$\{G\}_+ = \{A, B, C, D, E, F, G\}$  - 11  
 (pt. cō  $G \rightarrow EF$  și E este o cheie kandidat

$\Rightarrow$  Kandidatschlüssel: B, E, G

Sunt singurele cheie candidat de

c) Eine Relation R ist in der dritten

FDs F ist in der dritten Normal

für alle Abhängigkeiten  $A \rightarrow B$

1.  $B \subseteq A$  (FD ist trivial)
2. A enthält ein Schlüssel
3. B ist Teil von ein