

## „Jahrgangstest 2017/18“

- Gegeben sei eine Relation **Zehnkampf** mit dem Schema {SportlerName, Disziplin, ZeitHöheWeite, Punkte}

● **Z**

SN	D	ZHW	P
„Jim Knopf“	„100 Meter“	10,89	986
„Usain Bolt“	„100 Meter“	9,81	1723
„Jim Knopf“	„StabHoch“	4,56	976
...	...	...	...

- Welche funktionalen Abhängigkeiten gelten?
- Kandidaten-Schlüssel der Relation?
- In welcher Normalform befindet sich die Relation?
- Normalisieren Sie die Relation.
- Geben Sie in SQL die Sportler sortiert nach der Gesamtzahl ihrer Punkte aus.
- Wer hat Silber-Medaillen gewonnen?
- Wer sind die besten Sportler pro Disziplin? – in rel. Algebra

22

## „Jahrgangstest 2017/18“

- Fügen Sie in einen anfangs leeren B-Baum mit Knotenkapazität 4 die Datenelemente 10, 20, 30, 40, 50, 60, 90 ein. Wie sieht der Baum am Ende aus?
  - Löschen Sie jetzt die 20 und zeigen den resultierenden B-Baum.
- Betrachten Sie folgende/n Schedule/Historie zweier Transaktionen T1 und T2

w1(A) w2(B) r1(B) r1(C) r2(C) w1(C) c2 c1

- Ist diese Historie serialisierbar?
- Ist die Historie recoverable?
- Wie sieht der Serialisierbarkeitsgraph aus?
- Kann sie von einem strikten 2PL-Scheduler akzeptiert werden

23



# Elite-Studiengang SWT

## Datenbanksysteme

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

(1) In rel. Algebra soll auf der Relation

Zehnkampf: {[name, disziplin, höhezeit, punkte]}

ermittelt werden, welche Sportler in den jeweiligen Disziplinen am besten waren.

(2) im Tupelkalkül sollen die Sportler ermittelt werden, die in allen (in der Relation bekannten) Disziplinen angetreten sind.

Sie können die Abkürzungen ZK, N, D, HWZ, P verwenden.

1)  $R_1 \leftarrow \pi_{Z_1.N, Z_1.D} ( \rho(Zehnkampf) \mid X \mid \rho(Zehnkampf) )$   
 $Z_1.disziplin = Z_2.disziplin \wedge Z_1.punkte < Z_2.punkte$

~~$\pi_N (Z_N - R_1 \wedge \pi_N (Z_N - R_1))$~~

$\pi_{Z_1.N, Z_1.D} (Z_N) - R_1$

✓

2)  $\{ s.N \mid s \in ZN \wedge \forall st_1, st_2 \in ZN \wedge st_1.D \neq st_2.D \wedge \exists p \in ZN \mid p.D = st_1.D \wedge p.N = s.N \wedge p.D = st_2.D \}$

participation?

not equal



---

WS 2017/18  
Mo 10-14 Uhr

# Elite-Studiengang SWT

## Datenbanksysteme

---

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

---

### Quiz

Ermitteln Sie in SQL

1. welche Vorlesung hat die meisten direkten Voraussetzungen
2. In welchem Semester kann man "Der Wiener Kreis" frühestens hören (weil man vorher die direkten und indirekten Voraussetzungen hören muss)

# Elite-Studiengang SWT

## Datenbanksysteme

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

### Quiz

Erweitern Sie die Uni-Datenbank um die Relation

StudienPlan: {[Semester, VorlesungsNr]}

- Bestimmen Sie in rel. Algebra die Studenten, die alle für ihr Semester vorgesehenen Vorlesungen hören.
- Bestimmen sie in SQL die Studenten, die mindestens eine Vorlesung hören, die nicht für ihr Semester vorgesehen ist.

shortcut: OLP  
1. Obligatory lectures Pro Student  $\leftarrow \pi_{MatNr, VorNr} (Studenten \bowtie StudienPlan)$

Good Students hören Duplicate ~~MatNr~~  $\leftarrow \pi_{MatNr} (P(O \cap P) \bowtie P(O \cap P))$   
 $\begin{matrix} 01 \\ 02 \end{matrix}$   
 $01. MatNr = 02. MatNr$   
 $01. VorNr = 02. VorNr$

Duplicate Only MatNr  $\leftarrow \pi_{01. MatNr, 02. MatNr}$



---

# Elite-Studiengang SWT

## Datenbanksysteme

---

WS 2017/18  
Mo 10-14 Uhr

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

---

### Quiz

- (1) Sie erhalten (zumindest wenn Sie diese Aufgabe richtig lösen;-) einen einmaligen Bonus. Das heißt, ein (und nur ein) schwächstes **erzieltes** Quizergebnis wird für jede/n Studierende/n durch die **maximal** erzielbare Punktzahl ersetzt. Schreiben Sie den SQL Befehl für die Sicht QuizBonus.
- (2) Formulieren Sie (auf Deutsch oder Englisch) die Fragestellung der interessantesten bzw. schwierigsten der 100 SQL Aufgaben, die Sie diese Woche gelöst haben.
- (3) Die Relation Quiz ist nicht normalisiert. Warum nicht? Welche funktionale Abhängigkeit „ist schuld“ daran? Geben Sie den Schlüssel von Quiz an.

### Quiz

Qno	Name	Max	Erzielt
1	Meier	80	60
2	Meier	100	88
1	Schmidt	80	60
2	Schmidt	100	80

### QuizBonus

Qno	Name	Max	Erzielt
1	Meier	80	<b>80</b>
2	Meier	100	88
1	Schmidt	80	<b>80</b>
2	Schmidt	100	80



# Elite-Studiengang SWT

## Datenbanksysteme

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

### Quiz

Überprüfen Sie mittels einer SQL-Abfrage ob in der Relations-Ausprägung R mit dem Schema  $R=\{A,B,C\}$  die MVD

$A \twoheadrightarrow B$

eingehalten ist.

select \* from R r1, r2  
 where  $r1.a = r2.a$  and  
 $r1.b \neq r2.b$  and not  
 exists (select \* from R r3, r4  
 where  $r3.a = r4.a$  and  
 $r3.a = r2.a$  and  
 $r3.b = r1.b$  and  $r3.c = r2.c$   
 and  $r4.b = r2.b$  and  $r4.c = r1.c$ )

Wenn <sup>das</sup> Ergebniss nicht leer ist, dann ist die Ausprägung R nicht mvd erhaltend (Nonform)

### Zusatzfragen

(Allgemeinbildung eines DB'lers)

Wer erfand ... Ted Codd

... das relationale Modell

... den B-Baum Bayer

... den Rot/Schwarz-Baum Bayer

... das Transaktionskonzept

Wer gründete die erfolgreichste Datenbank-Firma Oracle founder Larry Ellison

Wer gründete Facebook Mark Zuckerberg

Nennen Sie mindestens zwei der vier SAP Gründer (wer von ihnen hat das Garching Planetarium finanziert?) !

Wie heißt SAP's Hauptspeicher-Datenbanksystem ?

Hana, Hana

Welche DB-Firma hat SAP übernommen?

Haupt DB Engine von MySQL

Was ist MariaDB und wem gehört MySQL?

Oracle [ MySQL nutzt 2 Engine  
MariaDB most common 2nd asom DB

Welche drei Datenbänkler gewannen den Turing-Award?



---

WS 2017/18  
Mo 10-14 Uhr

# Elite-Studiengang SWT

## Datenbanksysteme

---

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

---

### Quiz

Zeigen Sie den Aufbau eines R-Baums mit Knotenkapazität 4 mit folgenden Suchschlüsseln, die in der Gegebenen Reihenfolge einzufügen sind:

(1,20)(5,10)(10,5)(15,10)(10,15)(12,13)(13,5)(5,20)(3,20)(4,18)(1,10)(2,19)

Zeigen Sie wie die Bereichs-Suche (5..10,10..15) auf Ihrem Beispielbaum ausgeführt wird.

---

# Elite-Studiengang SWT

## Datenbanksysteme

---

---

WS 2017/18  
Mo 10-14 Uhr

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

---

### Quiz

Konzipieren Sie ein externes (also für über den Hauptspeicher hinausgehende Datengrößen) Sortiervverfahren, das besonders effizient ist bei fast sortierten Relationen, bei denen ein Tupel nie mehr als  $k$  Positionen "verrutscht" ist. Wie groß darf  $k$  bei Ihrem Verfahren maximal sein und welche Komplexität hat Ihr Verfahren?



WS 2017/18  
Mo 10-14 Uhr

# Elite-Studiengang SWT

## Datenbanksysteme

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

Sie fangen eine für Conny bestimmte verschlüsselte Nachricht mit dem Inhalt 2 ab. Conny's Zertifikat enthält den öffentlichen RSA-Schlüssel (11,35). Ermitteln Sie Conny's geheimen Schlüssel und entschlüsseln Sie die Nachricht.

$$N = 35, \text{ Prime Factorization } \Rightarrow 35 = 5 \cdot 7 = p \cdot q$$

$$\text{calculate totient function } \phi(n) = (p-1) \cdot (q-1) = (4) \cdot (6) = 24$$

Wir wissen schon  $e = 11$  aus der Zertifikat

$$\text{und } d \cdot e \equiv 1 \pmod{\phi(n)} \Rightarrow 11 \cdot d \equiv 1 \pmod{24}$$

using extended euclidian algorithm, we get the modular inverse

$$\Rightarrow d = 11 ; 11 \cdot 11 \equiv 1 \pmod{24}$$

$$\text{decrypted Message} = 2^d \pmod{35} = 2^{11} \pmod{35} = \boxed{18}$$

~~$$2^{11} = 2048$$~~



WS 2017/18  
Mo 10-14 Uhr

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

# Elite-Studiengang SWT

## Datenbanksysteme

### Quiz zur Serialisierbarkeit

(A) Geben Sie Beispiele von Historien (Schedules), die von einem strengen 2PL-Scheduler mit Deadlock-Erkennung mittels eines Wartegraphen akzeptiert würden aber **nicht**

1. bei der wound-wait Deadlockvermeidung bzw
2. bei der wait-die Deadlockvermeidung

„durchkommen“.

(B) Geben Sie ein Beispiel einer Historie, die beim strengen 2PL aber nicht bei der Zeitstempel-basierten Synchronisation entstehen kann.

(C) Geben Sie ein Beispiel einer Inkonsistenz, die bei gängigen Versionsverwaltungssystemen wie git (oder perforce, subversion, cvs, etc.) entstehen kann, weil diese Systeme auf snapshot isolation basieren. Warum wird in diesen Systemen keine „harte“ Serialisierbarkeit verwendet? Welches Versionsverwaltungssystem verwenden Sie persönlich für Ihre Projekte? Wie verhält es sich bei nicht möglicher Validierung gemäß der Snapshot Isolation? Was genau wird überprüft bei der Validierung?

Begründen Sie jeweils die Konstruktion Ihrer Beispiel-Historien und markieren Sie die entsprechenden Operationen in den Historien, die zur Differenzierung der Synchronisations-Verfahren führen.

Handwritten notes on the right margin:  
WS 2017/18  
Mo 10-14 Uhr  
Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr  
Quiz zur Serialisierbarkeit  
1. bei der wound-wait Deadlockvermeidung bzw  
2. bei der wait-die Deadlockvermeidung  
„durchkommen“.  
(B) Geben Sie ein Beispiel einer Historie, die beim strengen 2PL aber nicht bei der Zeitstempel-basierten Synchronisation entstehen kann.  
(C) Geben Sie ein Beispiel einer Inkonsistenz, die bei gängigen Versionsverwaltungssystemen wie git (oder perforce, subversion, cvs, etc.) entstehen kann, weil diese Systeme auf snapshot isolation basieren. Warum wird in diesen Systemen keine „harte“ Serialisierbarkeit verwendet? Welches Versionsverwaltungssystem verwenden Sie persönlich für Ihre Projekte? Wie verhält es sich bei nicht möglicher Validierung gemäß der Snapshot Isolation? Was genau wird überprüft bei der Validierung?  
Begründen Sie jeweils die Konstruktion Ihrer Beispiel-Historien und markieren Sie die entsprechenden Operationen in den Historien, die zur Differenzierung der Synchronisations-Verfahren führen.



# Quiz

22. Januar 2018

## 1 Top-k-Anfragen (80 Punkte)

Die in Abbildung 1 dargestellten Relationen Autos und Unterhalt dienen der Bewertung von Autos. Eine junge Studierende sucht ein Auto mit guter Balance zwischen Sportlichkeit und Kosten. Sie überlegt sich wie die drei Werte Preis, PS und monatlicher Unterhalt in einen Score umberechnet werden können und nutzt schließlich folgende Formel:

$$\text{Preis} - (100 * PS) + 24 * \text{Unterhalt}$$

Zeigen Sie die phasenweise Berechnung der Top-3 Ergebnisse jeweils mit dem Threshold- und dem NRA-Algorithmus.

Auto	Preis	PS	Auto	Unterhalt p. Monat
Seat Leon	25000€	200	Seat Leon	215€
Audi A1	17000€	96	Audi A1	220€
Citroen DS 4	20679€	100	Citroen DS 4	225€
Mini One	16500€	75	Mini One	262€
Mercedes C-Klasse	35000€	160	Mercedes C-Klasse	290€
Porsche Cayenne	80100€	420	Porsche Cayenne	430€

Abbildung 1: Relationen Autos und Unterhalt.

## 2 Radix-Bäume (20 Punkte)

Berechnen Sie die Anzahl der Knoten und den Speicherverbrauch eines *nicht* adaptiven Radix-Baumes mit Fanout 256 (Node256 des ART), wenn man  $2^{17} + 1$  32-bit Integer Werte (0 bis einschließlich  $2^{17}$ ) einfügt.

$$\frac{2^{17}}{2^8} * 2 \text{ Bytes} + 1$$



---

WS 2016/17  
Mo 10-14 Uhr

# Elite-Studiengang SWT

## Datenbanksysteme

---

Dozent: Prof. Alfons Kemper  
E-Mail: kemper@in.tum.de  
Telefon: +49 89 289-17254  
Büro: 2.11.54  
Bürozeiten: Di 13 Uhr

---

### Quiz

Zeigen Sie, dass ROWA und Majority/Consensus Spezialisierungen des Quorum/Consensus-Verfahrens sind, gemäß folgender Teilaufgaben:

16.17 Zeigen Sie („grob“), dass bei der write-all/read-any Methode zur Synchronisation bei replizierten Daten zwar nur serialisierbare Schedules erzeugt werden – unter der Voraussetzung, dass das strenge 2PL-Protokoll angewendet wird. Aber die Deadlock-„Gefahr“ nimmt dramatisch zu... warum?

16.18 Zeigen Sie, dass die write-all/read-any Methode zur Synchronisation replizierter Daten einen Spezialfall der Quorum-Consensus-Methode darstellt.

• Wie werden Stimmen zugeordnet, um write-all/read-any zu simulieren? • Wie müssen die Quoren  $Q_w$  und  $Q_r$  vergeben werden?

16.19 Einen weiteren Spezialfall des Quorum-Consensus-Verfahrens stellt das Majority-Consensus-Protokoll dar. Wie der Name andeutet, müssen Transaktionen sowohl für Lese- als auch für Schreiboperationen die Mehrzahl der Stimmen einsammeln. Zeigen Sie die Konfigurierung des Quorum-Consensus-Verfahrens für die Simulation dieses Majority-Consensus-Protokolls.