



Klausur zur Veranstaltung Einführung in Datenbanken (Sommersemester 2012)

Name: _____

Vorname: _____

Matrikelnummer: _____

Die Bearbeitungszeit beträgt 60 Minuten, sie können insgesamt 60 Punkte erreichen. Vor der Bearbeitungszeit haben Sie 5 Minuten Zeit, um die Klausur durchzulesen. **Sie dürfen mit der Bearbeitung der Klausur erst nach Ende dieser Einlesezeit beginnen, sobald dies angekündigt wird. Ein vorzeitiger Beginn wird als Täuschungsversuch gewertet.**

Bitte beachten Sie:

- Es sind keine Hilfsmittel zugelassen.
- Lösungen sind auf die gedruckten Linien zu schreiben. Sollte der Platz nicht ausreichen, befindet sich am Ende ein Ersatzblatt. Wenn auch dies verbraucht sein sollte, schreiben Sie bitte auf die Rückseiten.
- Bitte verwenden Sie keine Bleistifte oder Rotstifte für die Lösungen.
- Jeder Täuschungsversuch wird mit der Note 5,0 (nicht bestanden) bewertet.

Viel Erfolg!

Aufgabe 1: SQL

Gegeben sei das folgenden Schema. *Given the following schema.*

Tabelle Employee

<u>EmpNo</u>	FirstName	LastName	WorkDept	PhoneNo
10	Christine	Haas	A00	3978
20	Michael	Thompson	B02	3412
23	Eva	Pulaski	D04	5410
42	Masatoshi	Yoshimura	A01	3856

Tabelle Project

<u>ProjNo</u>	Name	RespEmp	Staff
154	Admin Services	23	4
326	User Education	20	12
378	Operation Support	20	6
410	Query Services	55	3

Tabelle Activity

<u>ActNo</u>	Key	Description
9541	Manage	Estimate Costs
9632	Code	Module Programming
9214	Test	Test Application
9630	Course	Develop new course

Tabelle EmpProjAct

<u>EmpNo</u>	ProjNo	ActNo	EmpTime
23	145	9214	1,0
23	326	9630	2,5
42	378	9214	0,5
10	326	9632	0,75

Erstellen Sie eine SQL-Abfrage für die folgenden Sachverhalte. *Create an SQL-Query for the following questions.*

a) Geben Sie alle Angestellten nach aufsteigender Abteilungsnummer sortiert aus.

List all employees. Order them by department number in ascending order.

```
SELECT * FROM Employee ORDER BY WorkDept ASC
```

- b) Geben Sie Nachname und Telefonnummer aller Angestellten sowie den Namen der von dieser Person verantworteten Projekte aus (RespEmp). Wichtig: geben Sie dabei auch die Angestellten aus, denen keine Projektverantwortung übertragen wurde.

List last name and phone number of all employees as well as the name of their responsibility taken projects (RespEmp). Important: also list the employees without responsibility for a certain project.

```
SELECT Employee.LastName, Employee.PhoneNo, Project.Name
FROM Employee LEFT JOIN Project ON Employee.EmpNo = Project.RespEmp
```

- c) Geben Sie die Abteilungsnummer der Angestellten (WorkDept), Projektname (Name), Aktivitätsschlüssel (Key) und benötigte Zeit (EmpTime) aller Einträge in der Datenbank aus. Nutzen Sie INNER JOIN.

List department numer (WorkDept), project name (name), activity key and time taken (EmpTime) for all DB entries. Use INNER JOIN.

```
SELECT e.WorkDept, p.Name, a.Key
FROM Employee e INNER JOIN EmpProjAct epa ON e.EmpNo = epa.EmpNo
      INNER JOIN Project p ON p.ProjNo = epa.ProjNo
      INNER JOIN Activity a ON a.ActNo = epa.ActNo
```

(Natürlich auch ohne Abkürzungen gültig!)

- d) Geben Sie die Beschreibung und die summierte Zeit (EmpTime) der Aktivität "Test" aus.

List the description and the total sum of time needed (EmpTime) for the activity "test".

```
SELECT s.Description, SUM(epa.EmpTime)
FROM EmpProjAct epa INNER JOIN Activity ON epa.ActNo = a.ActNo
WHERE a.Key = "Test"
```

(Natürlich auch ohne Abkürzungen gültig!)

Aufgabe 2: Normalisierung

- a) Durch Normalisierung werden Redundanzen in Datenbanken vermieden. Nennen Sie drei Gründe warum Redundanzen in Datenbanken verhindert werden sollten.

Normalization helps avoiding redundancy. Name three reasons why we should avoid redundancy in a database schema.

- Reduktion des Speicherplatzes
- Verhindern von Anomalien (ungewünschte Abhängigkeiten, Widerspruchsfreiheit, Konsistenz, Integrität)
- Klar strukturiertes Datenmodell
- ...

Folgende Aussagen sind entweder wahr oder falsch. Kreuzen Sie die entsprechende Antwort an und begründen Sie Ihre Aussage.

The following statements are either correct or wrong. Mark the correct column and explain your choice.

- b) Ein Relationenschema in erster Normalform mit dem Schlüssel, der nur aus einem Attribut besteht, ist in der zweiten Normalform.

A relational schema in first normal form that has a key consisting of only one attribute is in second normal form.

☐ Wahr / Correct ☐ Falsch/Wrong

Begründung/ Explanation:

Wahr; Nichtschlüsselattribute können nicht funktional von einem Teil des Schlüssels abhängig sein.

- c) Ein Relationenschema in zweiter Normalform, bei dem der Schlüssel aus nur einem Attribut besteht, ist in dritter Normalform.

A relational schema in second normal form that has a key consisting of only one attribute is in third normal form.

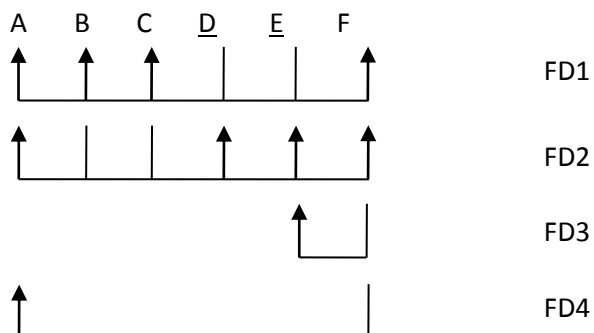
☐ Wahr / Correct ☐ Falsch/Wrong

Begründung/ Explanation:

Falsch; Es können immer noch transitive Abhängigkeiten bei den Nichtschlüsselattributen existieren.

d) Folgendes Schema ist gegeben. In welcher Normalform ist dieses Schema? Normalisieren Sie es in die nächst höhere Normalform.

Given the following schema, which normal form is it in? Normalise into the next higher normal form.



2. NF, da keine Attribute nur von einem Teil des zusammengesetzten Schlüssels (D, E) abhängig ist. Allerdings existiert noch eine transitive Abh. von einem Schlüssel (FD3). F ist entsprechend separat zu modellieren.

--

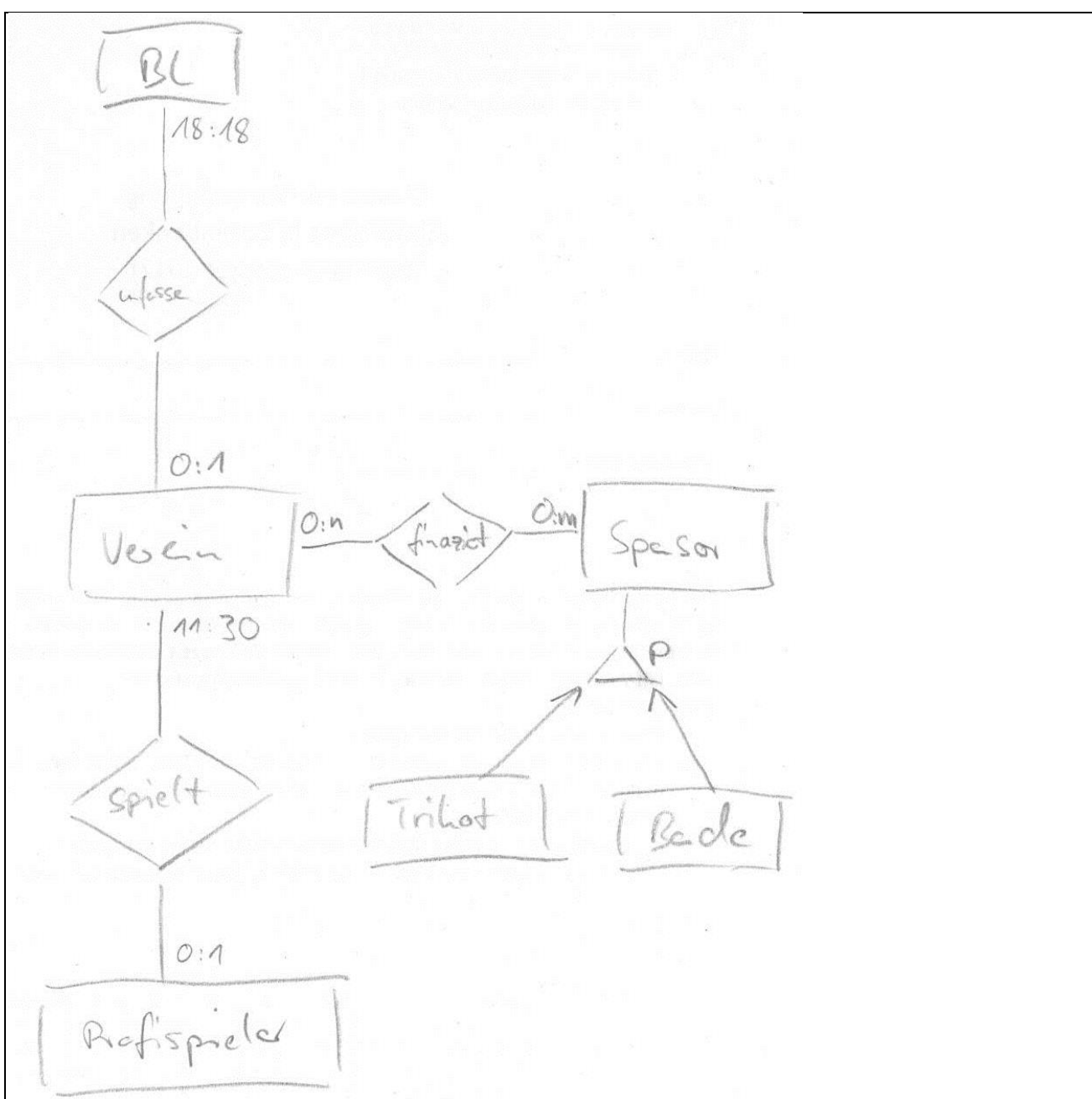
Aufgabe 3: Modellierung

a) Modellieren Sie folgenden Sachverhalt in einem ER-Modell und verwenden Sie dabei die min:max-Notation.

Create an ER-Model for the following scenario using the min:max notation.

Die Bundesliga umfasst 18 Fußballvereine, die bis zu 30 Profispieler unter Vertrag nehmen können. Ein Profispieler kann aber nur bei einem Verein beschäftigt sein. Zudem wird jeder Verein von Sponsoren finanziert. Hierbei werden zunächst Trikotsponsoren sowie Bandenwerbung unterschieden (zwar existieren weitere Sponsoren, diese werden allerdings nicht in diesem Sachverhalt modelliert).

The German premier league consists of 18 football (soccer) clubs, each of them can sign on up to 30 professional players. However, a player can only play for one club. Additionally, every club will be funded by sponsors. A distinction can be drawn between shirt sponsors and perimeter advertising (there might be other types of sponsors, however, we will not model them).



b) Erklären Sie den Unterschied zwischen einem mehrwertigen und einem zusammengesetzten Attribut.
Describe the difference between a multi-valued attribute and a composite attribute.

Mehrwertig: Attribut enthält mehrere Einträge (obere Schranke größer 1); Bsp. mehrere Tel.Nummern

Zusges: Hierarchie von Attributen; Bsp. Adresse teilt sich in versch. Attribute auf; Was davon letztendlich umgesetzt wird, wird erst im DV-Konzept entschieden.

Aufgabe 4: OLAP

Die Tabelle zeigt das (Teil-)Ergebnis einer SQL-Abfrage für die Verwendung in OLAP-Würfeln.
(Der Datenausschnitt ist willkürlich gewählt.)

The following table shows a (partial) result of an SQL-query optimized for OLAP-cubes.

(The data portion was chosen arbitrarily.)

Autoabsatz									
Modell	Motor	Absatzjahr	Quartal	Wirtschaftsraum	Land	Region	Rücklaufquote	Absatz	Umsatz
B3	2.0 TDI	2011	4	EMEA	DE	Stadt	0,22	15	54
B4	1.8 FSI	2011	3	EMEA	DE	Land	0,24	11	55
B3	1.6 FSI	2010	2	EMEA	EN	Stadt	0,33	14	23
B5	2.0 FSI	2011	4	EMEA	EN	Land	0,15	10	47
B3	1.6 FSI	2009	1	EMEA	FR	Stadt	0,30	7	66
B4	2.0 TDI	2010	4	EMEA	FR	Land	0,16	4	24
B6	2.5 TDI	2009	3	AMER	US	Stadt	0,22	22	30
B3	1.6 FSI	2011	2	AMER	US	Land	0,25	18	31
B8	4.0 TDI	2011	2	AMER	CA	Stadt	0,35	8	41
B4	1.6 FSI	2009	3	AMER	CA	Land	0,18	8	50
B4	2.0 TDI	2010	1	AMER	MEX	Stadt	0,21	12	55
B5	3.0 TDI	2010	4	AMER	MEX	Land	0,32	11	23
...

a) Unterscheiden Sie Dimensionen und Fakten im Würfel.

Name the dimensions and facts of the cube.

Fakten: Inhalt des Würfels; zu messende numerische Werte;

Dimensionen: Kategorisierungen der Messwerte; Spannen einen multidimensionalen Würfel auf; Können verschiedene Detaillierungsgrade bzw. Hierarchien aufweisen

b) Erstellen Sie eine OLAP-SQL-Abfrage, bei der die Rücklaufquoten aggregiert ausgegeben werden. Dabei soll die durchschnittliche Rücklaufquote für zwei Aggregationen ausgegeben werden: (1) nach Wirtschaftsraum, Land und Region; (2) nach Modell und Motor.



Create an OLAP-SQL-query that aggregates the average rate of return (Rücklaufquote). The result should contain two aggregations: (1) Economic area (Wirtschaftsraum), Country (Land) and Region; (2) Car Model (Modell) and Engine (Motor).

```
SELECT Wirtschaftsraum, Land, Region, Modell, Motor, AVG(Rücklaufquote)
FROM Autoabsatz
GROUP BY GROUPING SETS (Wirtschaftsraum, Land, Region), (Modell, Motor)
```

c) Nennen Sie drei Unterschiede in den Zielsetzungen von OLAP und Data Mining.

Name three differences in the objectives between OLAP and Data Mining.

OLAP	Data Mining
<ul style="list-style-type: none"> Hypothesis/ Verification driven „Does the revenue depend on the regions?“ 	<ul style="list-style-type: none"> Explorative/ Hypothesis driven „Which attributes of a customer do influence the buying behaviour?“
<ul style="list-style-type: none"> Shifting perspectives (slicing, dicing) 	<ul style="list-style-type: none"> Fixed objective of analysis
<ul style="list-style-type: none"> Highly interactive 	<ul style="list-style-type: none"> Partially autonomous
<ul style="list-style-type: none"> The relevance is determined by the research question 	<ul style="list-style-type: none"> The relevance of the results must be verified
<ul style="list-style-type: none"> Tool for analysis 	<ul style="list-style-type: none"> Tool for analysis and forecasting

  Datenmanagement - SoSe 2012 16

Aufgabe 5: XPath

Gegeben sei das folgende XML-Dokument. Geben Sie den XPath-Ausdruck für die folgenden Abfragen an:
Given the following XML document, write down the XPath expression for the following queries:

```
<?xml version="1.0"?>
<ipo>
  <shipTo export-code="1" type="EU-Address">
    <name>Helen Zoe</name>
    <street>47 Eden Street</street>
    <city>Cambridge</city>
    <postcode>126</postcode>
  </shipTo>
  <billTo type="US-Address">
    <name>Robert Smith</name>
    <street>8 Oak Avenue</street>
    <city>Old Town</city>
    <state>AK</state>
    <zip>95819</zip>
  </billTo>
</ipo>
```

```

<Items>
  <item partNum="833-AA">
    <productName>Lapis necklace</productName>
    <quantity>2</quantity>
    <price>99.95</price>
    <comment>Need this for the holidays!</comment>
    <shipDate>1999-12-05</shipDate>
  </item>
  <item partNum="748-OT">
    <productName>Diamond heart</productName>
    <quantity>1</quantity>
    <price>248.90</price>
    <comment>Valentine's day packaging.</comment>
    <shipDate>2000-02-14</shipDate>
  </item>
  <item partNum="783-KL">
    <productName>Uncut diamond</productName>
    <quantity>7</quantity>
    <price>79.90</price>
    <shipDate>2000-01-07</shipDate>
  </item>
  <item partNum="238-KK">
    <productName>Amber ring</productName>
    <quantity>3</quantity>
    <price>89.90</price>
    <comment>With no inclusions, please.</comment>
    <shipDate>2000-01-07</shipDate>
  </item>
  <item partNum="229-OB">
    <productName>Pearl necklace</productName>
    <quantity>1</quantity>
    <price>4879.00</price>
    <shipDate>1999-12-05</shipDate>
  </item>
  <item partNum="128-UL">
    <productName>Jade earring</productName>
    <quantity>5</quantity>
    <price>179.90</price>
    <shipDate>2000-02-14</shipDate>
  </item>
</Items>
</ipo>

```

a) Alle Elemente „item“.

All elements „item“.

```
//item
```

(Es geht natürlich auch immer eine absolute Adressierung!)

b) Das letzte Element „item“.

The last element „item“.

```
//item[last()]
```

c) Alle Kindelemente des Elements „item“ mit der partNum="229-OB".

All children of element „item“ with partNum="229-OB".

```
//item[@partNum='229-OB']/*
```

(Hinweis: ob ein oder zwei Slashes vor dem Stern sind in diesem Fall egal, da es keine weiteren Unterebenen gibt.)

d) Textinhalt des Elements „name“ mit dem Elternelement „billTo“.

Textual content of element „name“ with the parent element „billTo“.

Text(//billTo/name)

(Wichtig: relative Adressierung von //name geht in diesem Fall nicht)

Ersatzblatt / Additional Notes

[illegible]

