HOCHSCHULE FÜR TECHNIK STUTTGART
UNIVERSITY OF APPLIED SCIENCES

MASTER'S COURSE **S**oftware **T**echnology

EXAMINATION in **SUMMER SEMESTER 2013**

MODULE:	**Databases II**	NAME:	
DATE:	**9 July 2013**	SEMESTER:	
TIME:	**11.00 - 13.00 Uhr**		
EXAMINER:	**Prof. D. Koch**		

ALLOWED AIDS:	All materials that were distributed in class and in the Moodle, all your own notes, two text books of your choice, plus an English dictionary.
NOT ALLOWED:	Mobile Phones, laptop, and other communication devices
APPENDIX:	Schema and tuples of the universitydb database from the class notes.

Please write your name on each sheet that you turn in.

Turn in the problem sheets as well!

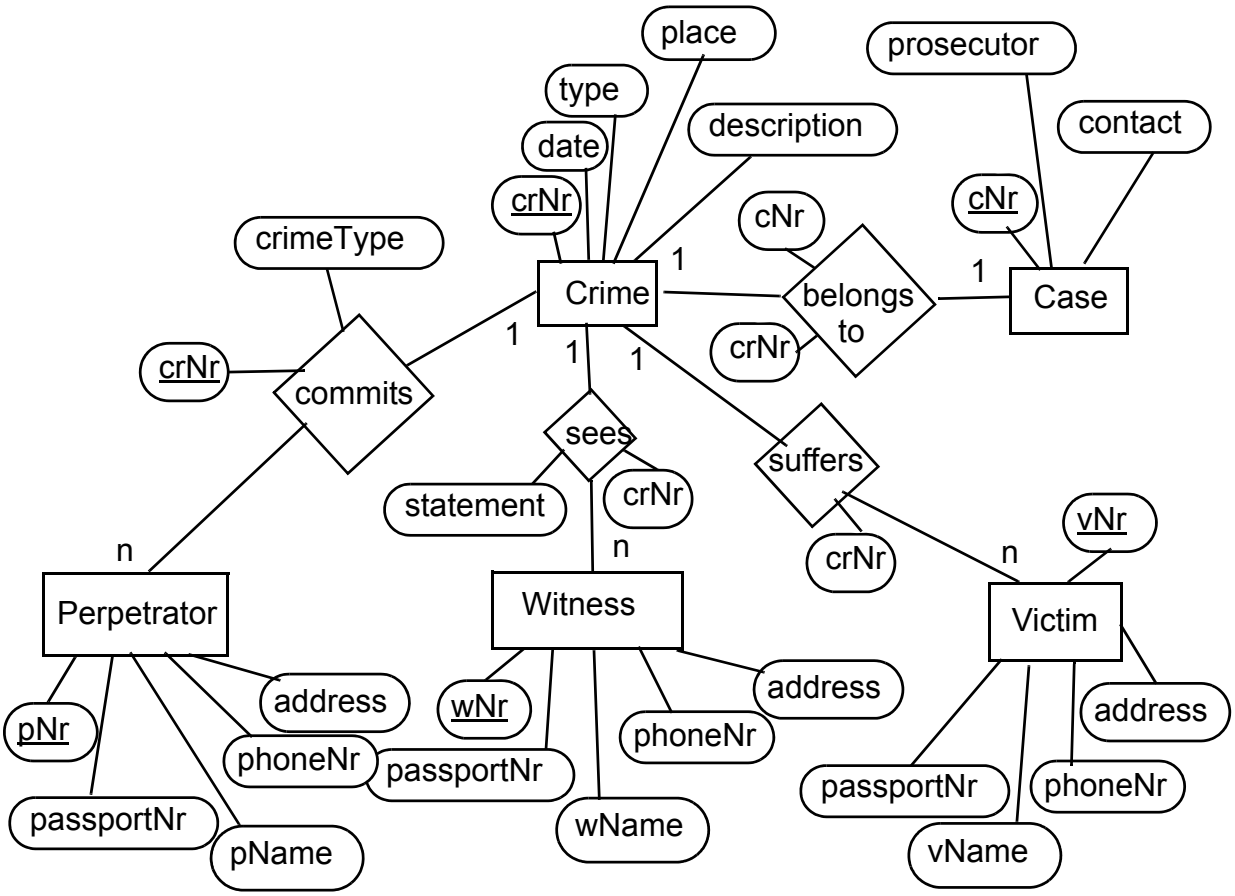
Problem	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	Σ
Maximum points	18	14	18	12	16	12	16	14	120
Achieved points									
Grade									

Problem 1. (3 times 6 points = 18 points)

A police officer with no database experience has created the following ER model that describes criminal offenses in a police database. The model is suboptimal in various respects which may concern form, structure and realism.
There are 6 types of problems in this model. You should find at least 3 of them.
I am NOT asking for additional entities, attributes, or relationships. There would be no limit to that. You may, however, restructure and rename the existing ones.

a) Draw a new, improved model.
b) Briefly explain which aspects were not good before and why they are better in your model.

(A few vocabulary explanations: perpetrator is someone who commits a crime. Witnesses give a statement about what they have seen. Prosecutor is someone who accuses a crime at a court of law).



Problem 2. (14 points)

Two application programmers, named A and B, try to program the same situation: A query is passed from a Java program to the universitydb we used in class, and the returned result is further processed in the Java program. In the following, you see part of the code of A and part of the code written by B.

a) (4 points)
What is the major difference between solution A and B? (Just name the concepts).

b) (10 points)
Briefly explain the relative advantages and disadvantages of A versus B in the context of this example.

Code of programmer A:

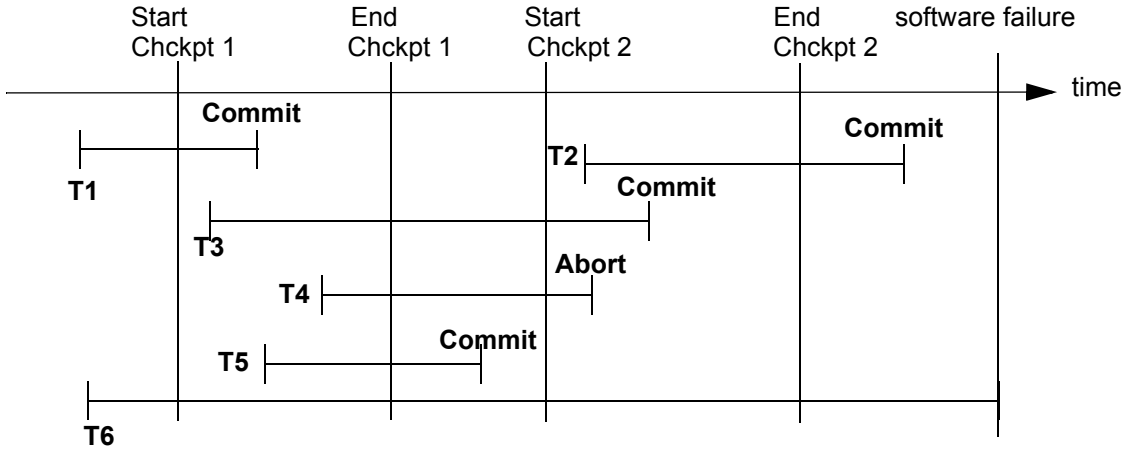
```
String classNr;
int matNr, grade;
String query = "select classNr, grade from Takes where matNr = 1234";
Statement stmt = conn.createStatement();
ResultSet result = stmt.executeQuery(query);
System.out.println ("classNr:  grade:");

while (result.next())
    classNr = result.getString(1);
    grade = result.getInt(2);
    System.out.println (classNr + ": " + grade);
}
```

Code of programmer B:

```
String query = KeyboardInput.readString(); //read query from user
Statement stmt = conn.createStatement();
ResultSet rs = stmt.executeQuery(query);
ResultSetMetaData rsmeta = rs.getMetaData();
int numCols = rsmeta.getColumnCount();

for (int i=1; i <= numCols; i++) {  // print the schema info
    System.out.println("Column #" + i + ": " +
                       rsmeta.getColumnName(i) + " of type "
                       + rsmeta.getColumnTypeName(i));
}// for
while (rs.next()) {
    for (int i=1; i<=numCols; i++){
        if (rsmeta.getColumnTypeName(i).equals("VARCHAR")){
            String value = rs.getString(i);
            System.out.print(value + "\t");
        } //if

        if (rsmeta.getColumnTypeName(i).equals("INT")){
            int value = rs.getInt(i);
            System.out.print(value + "\t");
        } //if
/**** Comment by Prof. Koch: imagine similar if-statements here for the
      other data types. I'm not printing them here for space reasons  ****/
            System.out.println();
    } //for
} //while
```

Problem 3. (3 times 6 points = 18 points)

Consider the following transaction execution scenario with a log file where fuzzy check-pointing is used:



a) (6 points)
List which transactions the recovery manager must undo, which ones it must redo, and for which ones no action is necessary, when the software failure interrups the system. Briefly explain your choice.

b) (6 points)
Now assume that Checkpoint 2 had **NOT** successfully finished. What is the answer of problem a) now? Briefly explain.

c) (6 points)
Assume again the original situation where checkpoint 2 has completed successfully. Now briefly explain the necessary actions of the recovery manager if the software failure happens in parallel with a crash of the databases disk.

Problem 4. (2 + 2 + 8 points = 12 points)

Consider the following schedule of two transactions A and B:
$A_{read\ (x)}$ $B_{read\ (y)}$ $A_{read\ (y)}$ $B_{read\ (x)}$ $A_{write\ (x)}$ $A_{read\ (z)}$ $B_{write\ (x)}$ $A_{write\ (y)}$ $B_{write\ (y)}$ $A_{write\ (z)}$

a) (2 points)
Is the schedule serial?
b) (2 points)
Is the schedule serializable?

c) (8 points)
If the schedule is serializable, explain why that is so.
If the schedule is not serializable, name the types of problems that occur,
and rearrange the operations to form a serializable, but not serial schedule of the transactions A and B.

Problem 5. (2 + 6 + 8 points = 16 points)

Consider two parallel transactions A and B in a system that uses the 2 phase locking protocol with record level locking. Transaction A uses isolation level read uncommitted. Transaction B uses serializable.
The data items X and Y are in different tables.
The following table shows the actions of the transactions in chronological order from top to bottom.

Transaction A (read uncommitted)	Transaction B (serializable)
Start transaction	
	Start transaction
	read x = 800
read y = 2000	
	write x:= x - 100 = 700 to the disk
read x = 700	
write y := y + x = 2000 + 700 = 2700 to disk	
	try to read y
	blocked
commit	
	read y = 2700
	rollback

a) (2 points)
State what the final values of X and Y are. (No explanation required here).
b) (6 points)
Briefly explain why the values of X and Y have these values at the end.
c) (8 points)
Are the values correct? Explain why. If something is incorrect, explain which transaction is guilty of the problem. If there is no problem, explain why a problem was avoided.

Problem 6. (12 points)

In comparison to the "immediate update strategy" and the "deferred update strategy" the "steal/no force strategy" suffers from the severe disadvantage that it is much more complex and requires a lot more implemented functionality for recovery. Name and explain the main reason why it is still the strategy that is most frequently used.

Problem 7. (4 times 4 points = 16 points)

Consider an online "Web Check-In Service" of an airline. Passengers who hold a flight ticket and a reservation code can check in 30 hours before take-off and book their seat in their flight. Alternatively, passengers can check in with a person at the airport.
A transaction typically consists of reading the list of available seats (they are usually displayed in a diagram of the plane) and making the reservation. It is possible to reserve several seats in one transaction in case of a group travelling together (expressed by the same reservation code).
It is essential that no two passengers will be checked in for the same seat.

Recommend the following parameters for the system:
a) Which transaction paradigm do you recommend for the execution of the transactions: ACID or BASE?
b) Should the transactions be nested, for instance if someone wants to reserve several seats for a group of passengers travelling together?
c) Should the system be distributed, i.e. should transactions run on a central server or on distributed sites?
d) Also state any disadvantages your system may have. Are they acceptable? Why?

Explain your decisions. You can answer all parts together, since they are related.
(This question requires a bit of discussion, something in the order of about 10-12 short sentences).

Problem 8. (14 points)

Consider the schema and the tuples in our universitydb (see appendix) which we used in the lab with MySQL.
Assume that the following code is now executed:

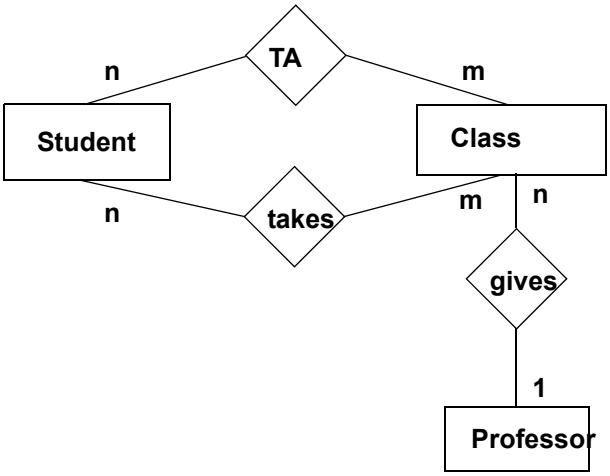
```
ALTER TABLE Student
ADD bestGrade double(3,1);

CREATE TRIGGER gpa_insert AFTER INSERT ON takes FOR EACH ROW
Update student
Set bestGrade = (SELECT MIN(grade) from takes where takes.matNr = NEW.matNr)
WHERE student.matNr = NEW.matNr;

CREATE TRIGGER gpa_update AFTER UPDATE ON takes FOR EACH ROW
Update student
Set bestGrade = (SELECT MIN(grade) from takes where takes.matNr = NEW.matNr)
WHERE student.matNr = NEW.matNr;

update takes set grade = 4.0 where classNr = 'DTB-SS93';
insert into takes values (2345, 'DTB-SS93', 2.0);
select * from student;
```

Your task:
Please state the output returned by MySQL after the last statement.

Appendix:
Our universitydb database in MySQL, installed in the lab

Consider the following ER model and relational model for a university database. For simplicity's sake, the attributes are omitted in the ERM.

Class			
classNr	**room**	**day**	**pName**
DTB-SS93	212	monday	Langes
BSY-SS93	114	monday	Müller
STP-WS92	212	tuesday	Müller
LIA-WS92	114	friday	Wagner



TA			
matNr	**classNr**	**hours**	**taSalary**
1000	LIA-WS92	40	1000
1234	LIA-WS92	35	1000
2345	DTB-SS93	80	2000
1000	STP-WS92	40	1000

Student	
matNr	**sName**
1234	Schmidt
2345	Schmidt
1000	Reinhard
4000	Wagner

Takes		
matNr	**classNr**	**grade**
1234	STP-WS92	1
1234	LIA-WS92	2
2345	STP-WS92	3
2345	LIA-WS92	4
1000	STP-WS92	3
1000	LIA-WS92	2
4000	STP-WS92	3
4000	LIA-WS92	1
1000	DTB-SS93	NULL
1000	BSY-SS93	NULL
1234	DTB-SS93	NULL

Professor	
pName.	**pSalary**
Langes	2000
Müller	3000
Wagner	2000
Schmidt	1000