

EXAMINATION in **SUMMER SEMESTER 2012**

MODULE: **Databases II**
DATE: **12 July 2012**
TIME: **13.30 - 15.30 Uhr**
EXAMINER: **Prof. D. Koch**

NAME:
SEMESTER:

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:

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	9	10	Σ
Maximum points	8	18	10	12	12	10	16	10	8	12	116
Achieved points											
Grade											

Problem 1. (8 points)

Briefly explain what is the major advantage of in-memory database systems as compared to traditional database systems? Explain how this advantage is achieved.

Problem 2. (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.

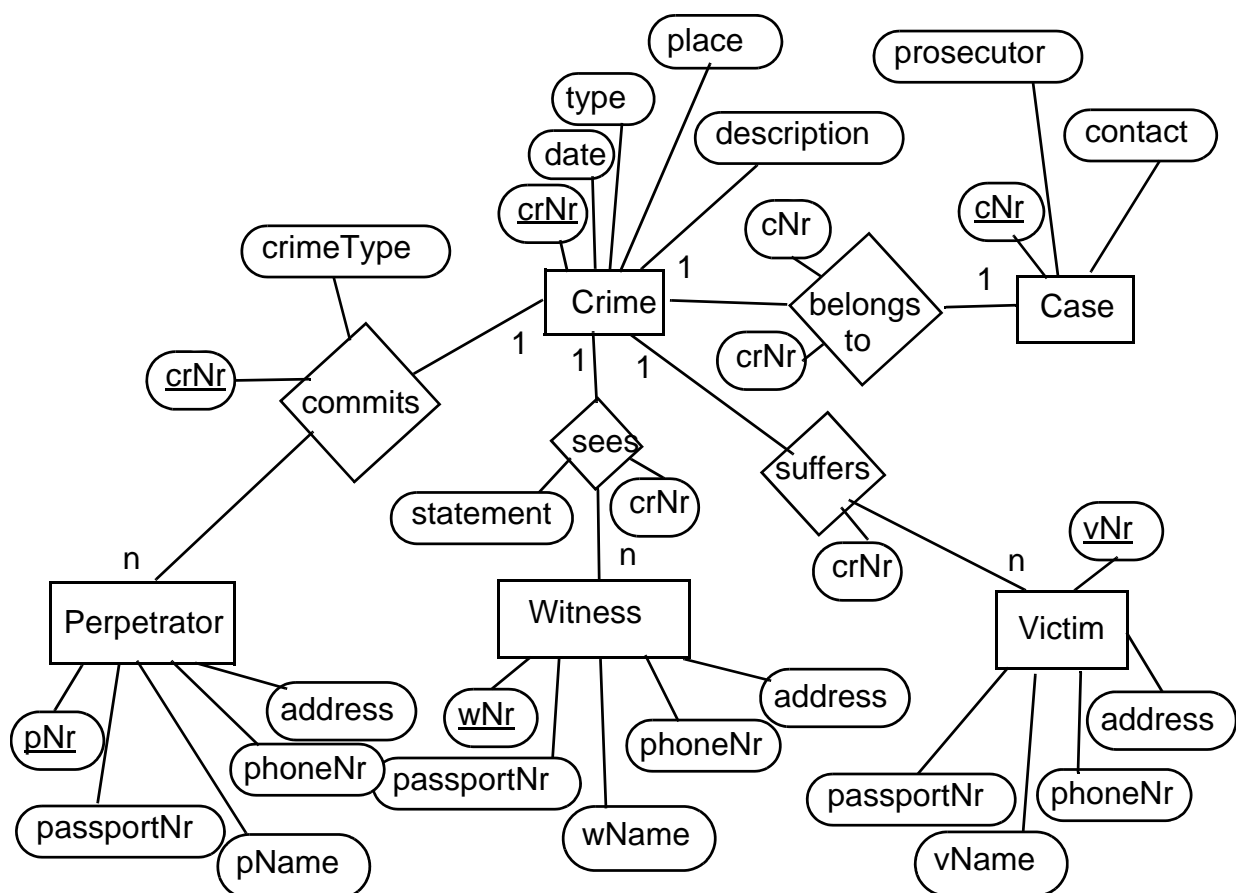
There are 6 types of problems in this model. You should find at least 3 of them.

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. Remember to underline primary keys.

b) Briefly explain which aspects were not good before and why they are better in your model.

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



Problem 3. (10 points)

Consider the following situation:

In a university's office of student affairs, the secretary needs a system that displays student information on the screen. The secretary always uses 10 different queries that have been well-known for years. It is not expected that she will ever need any other queries. Your task is to develop a new system to support this work, with a GUI and an application program in Java that uses JDBC to connect to a relational DBMS. Assume that for building the GUI, a software package is available that allows in a very simple way to generate the code for ad hoc queries and pass it as a String variable to the application program.

Question:

You now have the choice to

- a) implement the secretary's queries with the new GUI functionality and pass them as a String parameter to your application program, or
- b) to hard-code the queries in the application program.

State your decision and briefly explain your reasons.

Problem 4. (12 points)

Consider the relation Ancestors (pNr, personName, gender, birthYear, birthMonth, mother, father) which describes the ancestry of persons.

Mother and father are foreign keys referencing pNr. They describe for a given person which persons are their mother and father.

Write the following query in SQL:

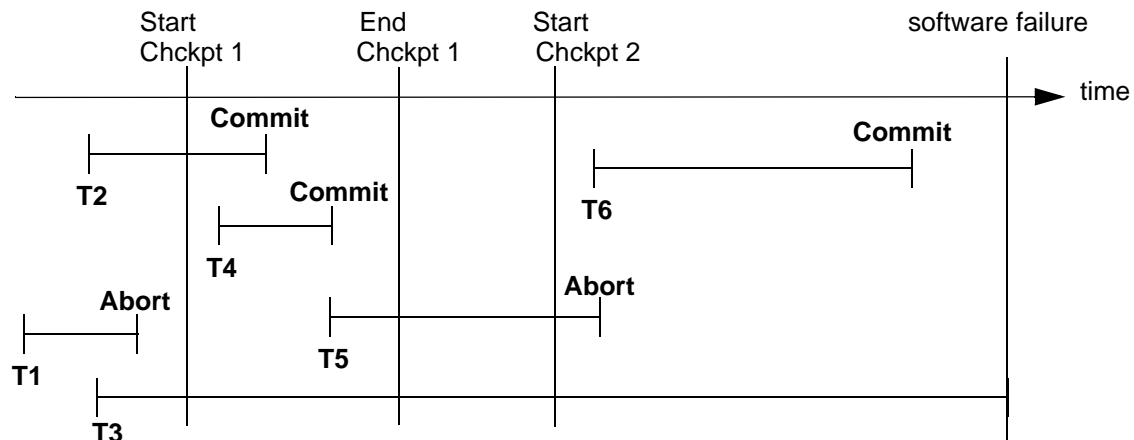
List the names and birthYears of all persons who are born in July and who are descendants in the female line of persons named Marie Curie¹.

"In the female line" means: at each level you look for all children, but for the next step further down you need to consider again only those who are mothers. (This is easier to write than the more general case).

1. Marie Curie was a female Polish-French physicist who discovered radioactivity and received Nobel prizes in physics and in chemistry. Her daughter also won the Nobel prize for chemistry.

Problem 5. (2 times 6 points = 12 points)

Consider the following scenario with a log file where fuzzy checkpointing is used:



a) (6 points)

List which transactions the recovery manager must undo and which ones it must redo. Briefly explain your choice.

b) (6 points)

Now assume that Checkpoint 2 had successfully finished just before the software failure occurred, after the Commit of T6. What is the answer of problem a) now? Briefly explain.

Problem 6. (2 + 2 + 6 points = 10 points)

Consider the following schedule of two transactions A and B:

$A_{\text{read}}(x) A_{\text{read}}(y) A_{\text{write}}(x) B_{\text{read}}(x) A_{\text{read}}(z) B_{\text{write}}(x) A_{\text{write}}(y) B_{\text{read}}(y) B_{\text{write}}(y) A_{\text{write}}(z)$

a) (2 points)

Is the schedule serial?

b) (2 points)

Is the schedule serializable?

c) (6 points)

If the schedule is serializable, explain why that is so.

If the schedule is not serializable, rearrange the operations to form a serializable, but not serial schedule.

Problem 7. (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 serializable. Transaction B uses read uncommitted.

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 (serializable)	Transaction B (read uncommitted)
Start transaction	
	Start transaction
read (x), x = 1000	
write $x := x - 200 = 800$ to the disk	
	read (y), y = 500
	read (x), x = 800
	write $y := y + x = 500 + 800 = 1300$ to disk
try to read (y)	
blocked	
	commit
read (y) = 1300	
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 8. (10 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 9. (8 points)

Briefly explain why there are **two** phases in the two phase commit protocol (2PC) for nested transactions. (You do not need to describe the whole protocol, and you will earn no points just for a description of the 2PC).

Problem 10. (12 points)

Consider the schema of the universitydb as used in class. Here, we need only the two tables TA (matNr, classNr, hours, taSalary) and Takes (matNr, classNr, grade).

Write a trigger in SQL to automatically execute the following behaviour:

When a new tuple is entered in Takes with a grade 1.0 and this student is a teaching assistant for this class, the teaching assistant's salary for this class is raised by 200.