group **A**

Please fill in your name and registration number (Matrikelnr.) **immediately**.

EXAM ON		06.12.2019
◯ DATENMODELLIERUNG (**184.685**)	◯ DATENBANKSYSTEME (**184.686**)	**GROUP A**
Matrikelnr.	Last Name	First Name

Duration: 60 minutes. Provide the solutions at the designated pages; solutions on additional sheets of paper are not graded. **Good Luck!**

Exercise 1: (6)

a) For the relational schemas (R, F_1) and (R, F_2), where $R = ABCDEFG$, find *all keys*.

FDs	Keys
$F_1 = \{BF \to EG, BD \to FG, AD \to BE, G \to AD\}$	....
$F_2 = \{BD \to C, BF \to A, CF \to DEG\}$	....

b) Consider the relational schemas (R, F_1) and (R, F_2), where $R = ABCDEFG$. Determine their corresponding normal forms, and mark the right answers.
Attention: for each correct solution: 1 point, for each wrong solution: -1 point, unanswered questions give 0 points. In total you get at least 0 points.

Dependencies	Keys
$F_1 = \{BCF \to G, BDG \to AE, CDF \to B, CDG \to B, CFG \to D\}$	`BCF , CDF , CFG`

neither 3NF nor BCNF ◯ 3NF & not BCNF ◯
BCNF & not 3NF ◯ 3NF & BCNF ◯

Dependencies	Keys
$F_2 = \{ABD \to C, ABF \to EG, AEF \to D, BCD \to F\}$	`ABD, ABF`

neither 3NF nor BCNF ◯ 3NF & not BCNF ◯
BCNF & not 3NF ◯ 3NF & BCNF ◯

Exercise 2: (9)

Assume a cinema storing information on its movies and visitors in a database with the following schema (primary keys are underlined, foreign keys are written in *italics*):

Movie(MovieName, Year, *GID: Genre.GID*)
Genre(GID, GenreName)
Visitor(VisitorID, Name)
watched(*VisitorID: Visitor.VisitorID*, *MovieName: Movie.MovieName*)
Review(*By: Visitor.VisitorID*, *MovieName: Movie.MovieName*, Stars)
(In the following, you may use suitable (unique) abbreviations for the names of both, relations and attributes.)

a) Consider the following query in the **relational algebra**. Briefly (**1 short sentence!**) describe the values returned by this query. (3 points)

$$\pi_{VisitorId}(\texttt{watched} \div \pi_{MovieName}(\texttt{Movie} \bowtie \sigma_{GenreName=\text{`Action'}}(\texttt{Genre})))$$

b) Consider the following query in the **domain calculus**. Determine the result of this query on the database instance given below. (3 points)

$$\{[vname, vid] \mid [vid, vname] \in \texttt{Visitor} \wedge \forall movie, b, s(([vid, movie] \in \texttt{watched} \wedge [b, movie, s] \in \texttt{Review}) \Rightarrow s = 5)\}$$

Database Instance:

Visitor	
VisitorId	**Name**
A	Alpha
B	Bravo
C	Charlie
D	Delta
E	Echo

watched	
VisitorId	**MovieName**
A	Playtime
A	Chungking Express
B	Ghost Dog
C	Waterworld
C	Chungking Express
D	Week-end

Review		
By	**MovieName**	**Stars**
A	Chungking Express	5
E	Chungking Express	5
B	Waterworld	2
C	Ghost Dog	3
B	Ghost Dog	5
C	Fallen Angels	5

c) We want to retrieve the names of all visitors who gave a 1 star review to at least one movie which they have not watched. Formulate a corresponding query in the **domain calculus**. (3 points)

Exercise 3: (8)

a) Consider the relational schema (R_d, F_d) where $R_d = ABCDEFG$ and F_d together with all keys are given below.

Using the (Relational) Synthesis Algorithm ("Synthesealgorithmus"), find a lossless and dependency preserving decomposition in 3NF (F_d is already in a canonical form/is a minimal cover).

For each schema R_i of the decomposition, state its attributes and mark one key by underlining.

$$F_d = \{G \to F, EG \to BD, EF \to G, BF \to G, AF \to C\}, \text{ Keys} = \{AEF, AEG\}$$

Decomposition into 3NF (Underline one key in each relation)

R1 .... R2 .... R3 .... R4 .... R5 ....

b) Consider the relational schema $R = (R_d, F_d)$ where $R_d = ABCD$ and $F_d = \{BC \to AD, BD \to AC\}$. For the decomposition $\{(R_1, F_1), (R_2, F_2)\}$ of R into two schemas with $R_1 = AC$ and $R_2 = ABD$, state whether this is a **lossless** decomposition.

If the decomposition **is lossless**, provide a short argument for your answer. (*Hint:* Just stating the definition of being lossless is *not* an argument.)

If the decomposition **is not lossless**, demonstrate this on a small example (i.e., provide a small instance for R_d and show that this instance contradicts the decomposition being lossless).

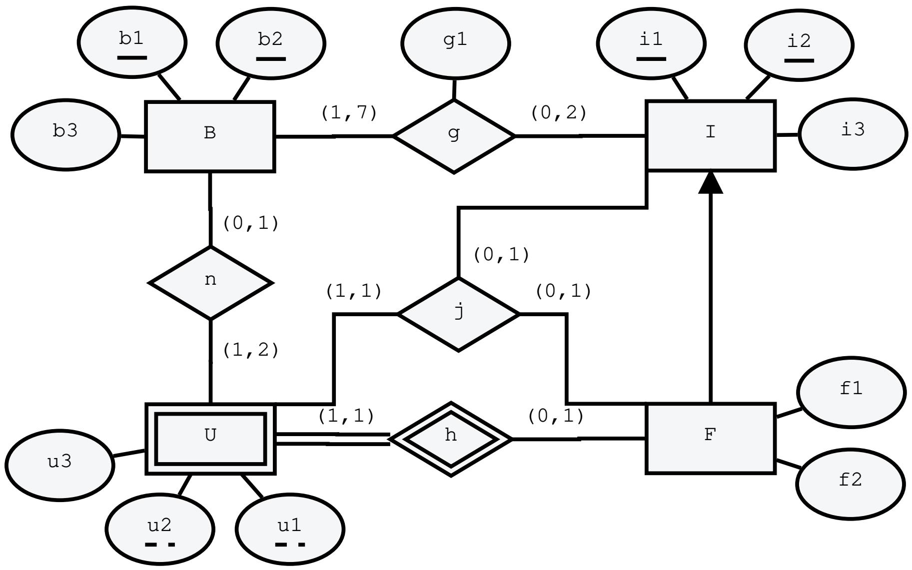
The decomposition is lossless: Yes ◯ No ◯

Justification:

(*Attention:* Just checking Yes/No without providing any justification gives no points!)

Exercise 4: (6)

Construct a relational schema according to the EER-diagram given below. For each relation, clearly mark the primary key by underlining the corresponding attributes. Mark foreign keys (FK) either by prefixing the name of the relation referenced by the FK (i.e., by `Relation.Attribute`) or by using the notation `NameOfAttribute`:`Relation.Attribute` (where `NameOfAttribute` is the name of the attribute in the current schema and `Relation.Attribute` describes the value that is referenced by the FK). You do not need to distinguish between FKs consisting of a single attribute and FKs combining several attributes.
Create as few relations as possible without introducing any redundancies. Note that the database does not allow `NULL`-values.



.... (....)

.... (....)

.... (....)

.... (....)

.... (....)

.... (....)

.... (....)

.... (....)

.... (....)

Exercise 5: (8)

Consider the relational schemas $R(\underline{A}BC)$, $S(\underline{B}D)$, and $T(\underline{SP}C)$.

Assume there exists an instance of R containing 4 tuples, an instance of S containing 3 tuples, and an instance of T containing 3 tuples. Thus

$$R(\underline{A}BC): 4 \qquad S(\underline{B}D): 3 \qquad T(\underline{SP}C): 3$$

Consider the expressions in Relational Algebra given below. For these expressions, provide the minimal and maximal possible size (= number of tuples) of their results over instances for R, S, and T of the given sizes. In addition, provide concrete instances over which the expressions actually realize these bounds, i.e. return results of minimal/maximal size. (Make sure that the provided instances contain exactly the given number of tuples.)

Attention: Points for correct instances are awarded only if the stated corresponding size is also correct!

a) **Expression:** $\pi_{A,B,C}(R \bowtie S) \cup (\rho_{C \leftarrow D} S \bowtie \rho_{A \leftarrow D} S)$

min. size of the result: . . . .

R		
A	**B**	**C**

S	
B	**D**

max. size of the result: . . . .

R		
A	**B**	**C**

S	
B	**D**

b) **Expression:** $\rho_{B \leftarrow P}(\rho_{A \leftarrow S}(T) \ltimes T) - \sigma_{A \neq 1}(R)$

min. size of the result: . . . .

R		
A	**B**	**C**

T		
S	**P**	**C**

max. size of the result: . . . .

R		
A	**B**	**C**

T		
S	**P**	**C**

Exercise 6: (8)

The instructions for this exercise are provided on the next page.

Overall: 45 points

You may separate this page form the exam and keep this page.

Thus, please do not provide any solutions on this page! Solutions written on this sheet will not be graded!

Instructions for Exercise 6:

Being asked too often for a pen, his colleagues have decided to provide Inspector Columbo an app as a replacement for his notebook as a Christmas present. You were chosen to develop the database for storing the sensitive notes and comments of the inspector. The database is supposed to store information on cases, witnesses, and their statements.

Create an EER-diagram based on the information described below. Use the (min,max) notation, and in case no explicit information is given, assume that there are no restrictions on the values for (min,max). The model shall work without using `NULL`-values, redundancies shall be avoided, and it is not allowed to introduce any attributes not described by the text. Finally, make sure that a key is defined for each entity type.

For each case, the identity of the victim (VICTIM) and the "Internal Police Case Number" (IPCN) must be stored. The identity of the victim and the IPCN are each unique for each case.

It shall further be possible to maintain information on each witness (INFO). Witnesses shall be identified by the combination of their Google Cookie ID (GCID) and their social security number (SSN). These two values shall be stored in the database as well. If witnesses are also suspects in a case, their alibi (ALIBI) and their motive (MOTIVE) in this case shall be stored. Witnesses can be suspects in several cases.

For each statement of a witness, the date (DATE) together with the actual statement (STMT) is to be stored. Each such witness statement is uniquely identified by the witness who gave the statement, the date of the statement, and an additional number. In addition, the inspector shall be able to mark statements that support other statements. Make sure that from the EER-diagram it becomes clear which statement is the supporting, and which the supported statement.

For meetings, the date (DATE), the place of the meeting (PLACE), and the number of police officers present (#P) have to be stored, with no two different meetings ever occurring at the same date at the same place. For each question, its content (CONTENT) is saved. There cannot be two different questions with the same content. If some question is a "final question", an additional flag (ACTUALLY) is used to indicate whether it is actually the final question.

It shall also be possible to record at which meeting, which question on which case was asked to which witnesses. At each meeting, at least one question on at least one case must be asked to at least one witness. Also, for each case there always exists at least one meeting where at least one question was asked to at least one witness.

In addition, each time a question is asked in such a setting, the reaction of the witness shall be stored (BEHAVIOUR).

Good Luck!