1. (5 points)BOOK (ISBN, Author, Title, Publisher, Year) abbreviated by BOOK (I, A, T, P, Y)

Suppose the following set *F* of functional dependencies are asserted to hold: TP→I, AP→T, I→ATP. Is the following decomposition of BOOK lossless? Why or why not? (A, T, P), (I, P, Y), (I, T)

2. You might explain your answer if you think it is needed, but no extra point will be given. (6 points)

Statement	TRUE	FALSE
Serial schedules are serializable		
A schedule is view serializable in the case we can get a serial schedule by swapping no conflicting instructions		
For 4NF we have also to consider multivalued dependencies for finding the candidate keys.		
In a deferred modification technique the transaction is committed if the log record <T, COMMIT> has been written into the log file.		
Two phase locking guarantees deadlock prevention, but does not ensure conflict serializability.		
If A→B holds then so does A→→B		
In a deferred modification technique the transaction is committed if the log record <T, COMMIT> has been written into the log file and the log file has been saved onto a nonvolatile storage.		
When the timestamp-based concurrency control is applied, every schedule is view equivalent to a serial schedule with respect to their timestamps.		
IF a relation is in BCNF, it is also in 3NF		
If a relation is in BCNF, it is also in 4NF		
A relation is in BCNF, if it is in 3NF and prime attributes should not depend on any other attribute set except for superkeys.		
A relation is in 3NF if it is in 2NF and nonprime attributes do not depend on any candidate key.		

4. (5 points)
4.1 Give 2 definitions for 3NF.
Definition1:

Definition2:

4.2 Prove only one of those: Definition1 implies Definition2
Definition2 implies Definition1

Which direction will you prove, write here (2 points):

4.3 Proof:

5. Explain why the algorithm for lossless decomposition into BCNF does ensure losslessness.

5.1 Write here the definition of a lossless decomposition:
5.2 Write here the algorithm:
5.3 Write here your detailed explanation for losslessness:

6.(10 points)

Suppose we have a database for an investment firm, consisting of the following attributes:

DB (Broker, Office_of_broker, Investor, Stock, Quantity, Dividend)

Quantity=quantity of stock owned by an investor

Dividend=dividend paid by a stock

Hence the overall schema is R(B,O,I,S,Q,D). Assume the following functional dependencies are required to hold on this database:

$I \rightarrow B$, $IS \rightarrow Q$, $B \rightarrow O$, $S \rightarrow D$

6.1 List all candidate keys for *R*.

6.2 Consider the following database instance D_1 of *R*:

B	*O*	*I*	*S*	*Q*	*D*
Merrill	SLC	Greene	IBM	100	$1.50
Schwab	Provo	Hatch	Unisys	200	$0.70
Edwards	Loa	Orton	Novell	300	$0.05
Carl	Bicknell	Hansen	Borland	400	$2.00
Schwab	Provo	Hatch	Novell	500	$0.10

Is D_1 consistent with the dependencies specified above? Why or why not?

6.3

Give a lossless decomposition of *R* into Boyce-Codd Normal Form.

6.4 Does your answer to Question 8.3 preserve all given and implied functional dependencies? Explain.

6.5 Give a lossless decomposition of *R* into Third Normal Form, preserving functional dependencies.

6.6 Is your answer to Question 8.5 in BCNF? Your answer can be simply **Yes** or **No**, without further explanation

7. (8 points)

Step	T_1	T_2	T_3	T_4
1		READ(W)		
2	READ(X)			
3	WRITE(X)			
4				READ(V)
5	WRITE(Y)			
6			READ(Y)	
7				WRITE(X)
8			READ(V)	
9		WRITE(W)		
10		READ(Y)		
11			WRITE(X)	
12			WRITE(Y)	
13	READ(Z)			

7.1 Is this a conflict serializable schedule? If **yes**, show an equivalent serial schedule for $T_!$, T_2, T_3, T_4. Explain. If **no**, argue why not. Hint: Draw the precedence graph! (hints will be not given in the midterm! ☹

Similar: a precedence graph is given, is the schedule conflict serializable?

7.2 Add LOCK-X(), LOCK-S() and UNLOCK() statements to transaction T1 and T2 satisfying the two-phase protocol. Sign and name the 2 different phase.
Can you refine the schedule introducing upgrades and downgrades?

7. 3. Suppose that in the schedule above we apply timestamp protocol.

Transaction	Timestamp
T1	10
T2	20
T3	30
T4	40

Which of the transactions has to be aborted?

8.(6 points)

Consider a database management system running transactions concurrently. The table below shows a part of the log-file. What type of logging technique is used?

9.1 Explain, what is the difference in log records using the other technique we learnt.

Step	*Log Entry*
1	<T_1 **start**>
2	<T_1, *A*, 100, 200>
3	<T_2 **start**>
4	<T_2, *C*, 700, 800>
5	<T_3 **start**>
6	<T_2, *D*, 900, 300>
7	<T_3, *E*, 9, 30>
8	<T_2 **commit**>
9	< **checkpoint** *S* >
10	<T_3, *F*, 15, 20>
11	<T_3 **commit**>
12	<T_1, *H*, 75, 50>
13	<T_4 **start**>
14	<T_4, *G*, 57, 25>
15	<T_4 **commit**>
16	<T_1, *B*, 104, 204>
17	<T_1 **commit**>
18	<T_5 **start**>
19	<T_5, *I*, 33, 44>
20	<T_5, *C*, 800, 850>
21	<T_6 **start**>
22	<T_6, *D*, 300, 350>
23	<T_6 **commit**>

24	<T_5 **abort**>

9.2 In log record 9, what should be the value of list S ?

9.3 Suppose the log ends at line 19 (inclusive), and a power failure occurs. What action has to be taken and in what order during recovery?

9.4 Suppose the log ends at line 16 (inclusive), and a power failure occurs. What action has to be taken and in what order during recovery if we apply the other modification technique?

9.5 Suppose the log ends at line 9 (inclusive), and a power failure occurs. Show the values of all database items after recovery, supposing the original modification technique concluded from the log.

A	B	C	D	E	F	G	H	I

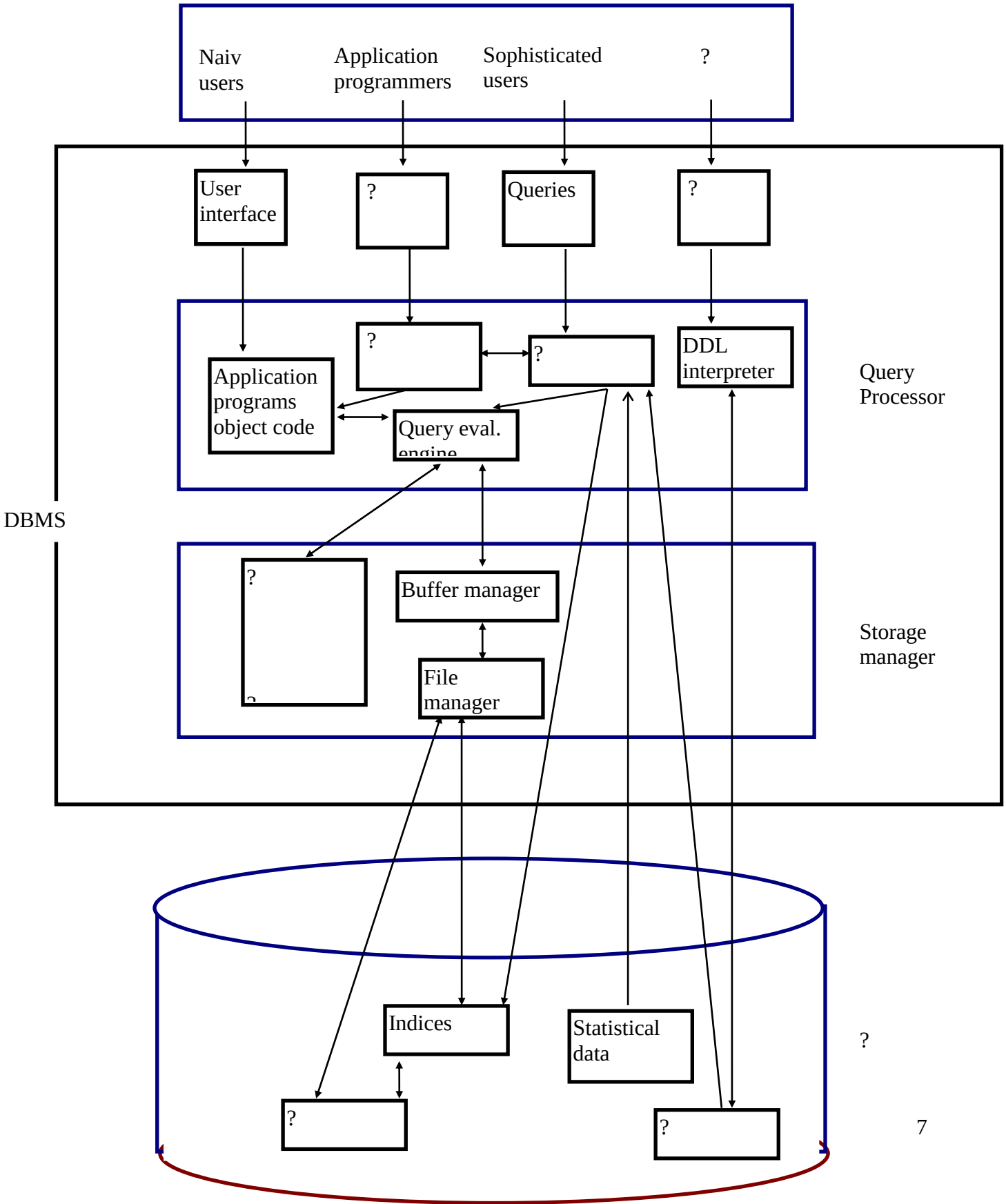
10. What do you know about embedded SQL?

11. What do you know about triggers?

12. Explain the program below… embedded SQL program…

13. Fill the gaps (denoted by ?) in the figure next page. Explain the tasks of these units.

.



13. a.) What are the levels of data abstraction?
b.) Explain what each level means.EZ BIZTOSAN NEM LESZ, csak majd a vizsgán ☺

14. a.) What are ACID properties?

A stands for….
Meaning:

C stands for….
Meaning:

I stands for….
Meaning:

D stands for…
Meaning:

b.) Which part of the DBMS is responsible for

A_____ C_____

I_____ D_____

15.
What is shadow paging recovery?
16.
What is timestamp-based protocol?

18.. Consider the following database log records.
a.) Decide the mode of the database modification recovery scheme (ie. deferred or immediate)
b.) What operations will be done and what will be the values of A, B, C if the system chrashes:
<u>Between f, g:</u> operations will be done in this order in case log file (i):
values in the database case (i) before the crash: A= B= C=

<u>Between j, k:</u> operations in will be done in this order in case log file (i):

values in the database case (i) before the crash: A= B= C=

<u>Between f, g:</u> operations will be done in this order in case log file (ii):

values in the database case (ii) before the crash: A= B= C=

Between j, k: operations will be done in this order in case log file (ii):

values in the database case (ii) before the crash: A= B= C=

(i) mode (deferred or immediate)? (ii) mode(deferred or immediate)?

a.) *<T1, start>* a.)<T1, start>
b.) *<T2, start>* b.)<T2, start>
c.) *<T1, A, 0, 1>* c.)<T1, A, 1>
d.) <T2, C, 30, 33> d.)<T2, C, 33>
e.) <T3, start> e.)<T3, start>
f.) *<T1, B, 20, 22>* f.)<T1, B, 22>
g.) *<T1, commit) >* g.)<T1, commit>
h.) *<T3, B, 22, 222>* h.)<T3, B, 222>
i.) *<T2, A, 1, 11>* i.) <T2, A, 11>
j.) <T2, commit)> j.) <T2, commit>
k.) *<T3, A, 11, 111>* k.) <T3, A, 111>
l.) *<T3, commit>* l.) <T3, commit>

19 .Consider the following transactions:

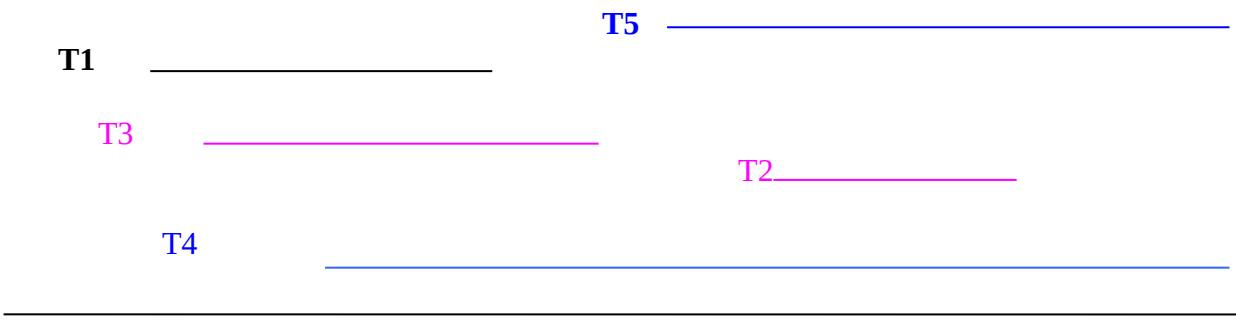
T1:	T2:
read (A)	read (B)
read(B)	read(A)
if A=0 then B:=B+1	if B=0 then A: = A+1
write(B)	write(A)

a.) Add lock and unlock instructions to transactions T1 and T2 so they observe the two-phase locking protocol.
b.) Give an appropriate schedule (different from the serial one).
c.) What are the names of the two phases? Draw each phase in your schedule.

20.
Suppose that transactions are run in the system modeled by the figure below. Explain what would be the recovery process in case of

a.) deferred
b.) immediate modification technique.

Write down precisely what actions have to be carried out, and in which order? (Use the appropriate words, for example RESART#REDO.)



time

CHECKPOINT FAILOR

21. Consider the attributes R with A B C D E F G which have the following functional dependencies:
Fd={AD → F, AE → G , DF → BC, E → C, G → E}
Consider the decomposition into 3 relations: (ADF) (EC) (ABDEG).

This decomposition (IS) (IS NOT) lossless, because…… (4 points)

If the relation is not in BCNF then give a lossless BCNF decomposition. Use the algorithm we learnt.

What is the best normal form for R?

If it is not in BCNF then give a lossless BCNF decomposition. Is your BCNF decomposioton lossless?

If it Is not in 3NF give a lossless dependency preserving 3NF decomposition.

22. Would the join of the three relations corresponding to the following relational schemas:

R1(A, B, G) R2(C, E) R3(A, C, D)

be lossless? To answer this question, consider the following set of associated functional dependencies:

F= {C -> E, B -> G, A -> B, D -> B}

23. 4. R(A, B, C, D, E, G) F:= {C -> E, B -> G, A -> B, D -> B}

(1) Find CK for R!

(2) Give at least 3 superkeys !
(3) What is the best normal form for R?
(4) If it is not in 2NF, give a lossless 2NF (BCNF) de composition.

24.
Given relation R(A, B, C, D, E, G, H, I), and its minimal cover :
Minimal cover: {H→G, H→C, H→E, E→D, BD→A})
Give a lossless dependency preserving 3NF decomposition using canonical or minimal cover (only one of them! :). Show the steps we learnt. (2 points). Check using tabular method that your decomposition is really lossless. (2 points) Is this decomposition in BCNF as well?

25.
4. Consider the attributes R with A B C D E F G which have the following functional dependencies:
Fd={AD → F, AE → G , DF → BC, E → C, G → E}

25. 1 List all candidate keys, explain:
(2 point)

25. 2. Give 3 superkeys:
(1 point)

25.3 Give a primary key:
(1 point)

25.4 Which of the following dependencies are implied by those in ***Fd*** above? If your answer is YES, please explain. If you can explain in 2 different ways, then you might earn extra points. (1 point each)

a. ADC → B : YES/NO, because....

b. A → G: YES/NO, because....

25.5 Give 2 elements of Fd^+ being different from the trivial dependencies. (2 point)

25.6 Find F-. Based on this, give a lossless dependency preserving 3NF decomposition

25.7 Consider the attributes R with A B C D E F G which have the following functional dependencies:
Fd={AD → F, AE → G , DF → BC, E → C, G → E}
Consider the decomposition into 3 relations: (ADF) (EC) (ABDEG).

This decomposition (IS) (IS NOT) lossless, because…… (4 points)

25.7 If the relation is not in BCNF then give a lossless BCNF decomposition. Use the algorithm we learnt.

26. Questions:

Why normalization is needed?
What are anomalies?
What are abstraction levels?
What is 2 phase protocol?
What is shadow paging?
What is logging?

What are Armstrong axioms?
What is the role of attribute closure, dependency set closure?
Give the definition of a CK, SK, PK.
Write down the following algorithm:

- **attribute closure**
- **getting SK**
- **lossless BCNF decomposition**
- **lossless AND dependency preserving 3NF decomposition**
- **recovery processes**

PL-SQL: a light question, you surely know… Ask Dr. Tam….