Model for the Written Exam

Computer Science in English, 2nd year - Database Management Systems

I. 1. Define a *serializable schedule*. **2p**

2. <u>Briefly</u> describe the *voting phase* of the Two-Phase Commit protocol.

II. Solve the following problems: **4p**

1. Consider the schedule below: (0.5p)

T1	**T2**
read(A)	
A = A + 50	
write(A)	
	read(A)
	A = A * 4
	write(A)
	read(B)
	B = B * 4
	write(B)
read(B)	
B = B + 50	
write(B)	

time

The following integrity constraint must hold: *A = B.* Before the execution of T1 and T2, A = 100 and B = 100. Is the schedule serializable? Justify your answer.

2. Consider the following schedule over transactions T1, T2, T3, T4 (all transactions commit): (1p)

T1	**T2**	**T3**	**T4**
W(A)			R(C)
	R(B)		
		W(D)	
	R(A)		
R(D)			
			W(B)
R(C)			

time

a. Draw the precedence graph.

b. Is the schedule conflict serializable? Justify your answer.

3. Express the SQL query below in the extended relational algebra, using a $\sigma\ \pi \times$ expression. (1p)

```
SELECT R1.IDR1, MAX(R2.C2)
FROM Relation1 R1, Relation2 R2, Relation3 R3
WHERE R1.IDR1 = R2.IDR1 AND R2.IDR2 = R3.IDR2 AND R3.Title = 'TBA' AND R1.C1 = 20
GROUP BY R1.IDR1
```

4. Let R and S be 2 relations. R has 10.000 records; a page can hold 10 R records. S has 2.000 records; a page can hold 10 S records. (1.5p)

a. 52 buffer pages are available. Compute the cost of $R \otimes_{R.a=S.b} S$ using *page-oriented nested loops join* and *block nested loops join*; S is the outer relation.

b. Compute the cost of sorting R using *external merge sort* with 200 buffer pages.

c. R is stored at Madrid, S is stored at Athens. Compute the cost of $R \otimes_{R.a=S.b} S$ using *simple nested loops join (tuple-oriented)* in Athens, without caching; S is the outer relation.

III. Choose the correct answer(s) for the following 14 multiple choice questions. Each question has at least one correct answer. Fill in the encoded data for question 15. Enter all answers in the table below. **3p**

1.	2.	3.	4.
5.	6.	7.	8.
9.	10.	11.	12.
13.	14.		
15.			

1. T1 and T2 are 2 concurrent transactions, both active at time *t*.

a. The following execution describes a *write read* conflict: At time *t*, T2 is reading a data object previously written by T1.

b. The following execution describes a *write read* conflict: At time *t*, T2 is writing a data object previously read by T1.

c. The following execution describes a *read write* conflict: At time *t*, T2 is reading a data object previously written by T1.

d. The following execution describes a *read write* conflict: At time *t*, T2 is writing a data object previously read by T1.

e. none of the above answers is correct.

2. A schedule S:

a. is conflict serializable if and only if its precedence graph has exactly one cycle.

b. is conflict serializable if and only if its precedence graph is acyclic.

c. is conflict serializable if and only if its precedence graph has exactly two cycles.

d. is conflict serializable if and only if its precedence graph has exactly three cycles.

e. none of the above answers is correct.

3. Under the READ UNCOMMITTED isolation level:

a. S locks must be acquired to perform read operations.

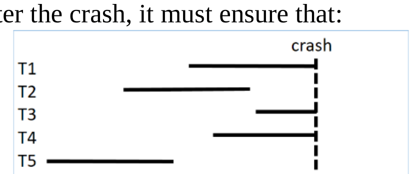
b. read operations are performed without acquiring S locks.

c. X locks must be acquired to perform write operations.

d. write operations are performed without acquiring X locks.

e. none of the above answers is correct.

4. Consider the execution below. When the system comes back up after the crash, it must ensure that:



a. T1, T3, T4 are durable; T2 and T5 are undone.
b. T1, T3, T4 are undone; T2 and T5 are durable.
c. T1 is undone only if T2 and T4 are also undone.
d. T2 is durable only if T5 is undone.
e. none of the above answers is correct.

5. When using a *steal*, *no-force* strategy:
a. the changes of a transaction in the buffer pool can be written to disk before the transaction commits.
b. the changes of a transaction in the buffer pool cannot be written to disk before the transaction commits.
c. when a transaction commits, all its changes in the buffer pool must be immediately written to disk.
d. transactions are not required to be atomic and durable.
e. none of the above answers is correct.

6. In horizontal fragmentation:
a. the reconstruction operator is the natural join.
b. the union of the horizontal fragments must be equal to the original relation.
c. fragmentation is performed with projection operators.
d. fragmentation is performed with selection predicates.
e. none of the above answers is correct.

7. In the Two-Phase Commit protocol:
a. if at least one subordinate votes *no*, the corresponding transaction is aborted.
b. a transaction can be aborted only if all subordinates vote *no.*
c. transactions are always committed.
d. if at least one subordinate votes *yes*, the corresponding transaction is committed.
e. none of the above answers is correct.

8. In data replication:
a. *primary site replication* is an asynchronous replication technique.
b. *primary site replication* is a synchronous replication technique.
c. *read-any write-all* is a synchronous replication technique.
d. *read-any write-all* is an asynchronous replication technique.
e. none of the above answers is correct.

9. Let T[<u>TID</u>, A, B, C] be a relation. Tuples in T with A < 70 are stored at Madrid; tuples with A >= 70 are stored at Athens. T is not replicated. Consider query Q below:

```
SELECT B, C
FROM T
WHERE A BETWEEN 50 AND 100
```

a. The DBMS executes Q at Madrid and Athens, then takes the union of the obtained results.
b. Q is executed only at Madrid.
c. Q is executed only at Athens.
d. Tuples in Athens must be brought to Madrid; Q must be executed at Madrid.
e. none of the above answers is correct.

10. A database access request contains:
a. the requesting user.
b. the criminal record of the requesting user.
c. the operation the user wants to perform.
d. the requested object.
e. none of the above answers is correct.

11. I is an index with search key <C1, C2, C3, C4>.
a. If I is a hash index, I matches condition *C1 > 10 AND C2 > 7*.
b. If I is a hash index, I matches condition *C1 = 10 AND C2 = 7 AND C3 = 1 AND C4 = 5*.
c. If I is a B+ tree index, I matches condition *C1 = 10 AND C2 = 7*.
d. If I is a B+ tree index, I matches condition *C2 = 7 AND C3 = 9*.
e. none of the above answers is correct.

12. Let R be a relation with P pages. The cost of sorting R using *simple two-way merge sort* (i.e., with 3 pages in the buffer pool) is:
a. π^{P}
b. $2P(\lceil \log_4 P \rceil + 1)$
c. $2P(\lceil \log_2 P \rceil + 1)$
d. $2P(\lceil \log_3 P \rceil + 1)$
e. none of the above answers is correct.

13. In a *left-deep* tree:
a. both children of each join node are base relations.
b. the left child of each join node is a base relation.
c. there are no joins.
d. the right child of each join node is a base relation.
e. none of the above answers is correct.

14. Consider the query:

```
SELECT *
FROM R1, R2, R3
WHERE p1 AND p2 AND p3
```

The conditions tested by the predicates in the WHERE clause are statistically independent. The cardinality of a relation R is denoted by |R|. The reduction factor associated with predicate p is denoted by RF(p).
The cardinality of the query's result set can be estimated by:
a. $\frac{|R1|*|R2|*|R3|}{RF(p1)+RF(p2)+RF(P3)}$
b. $|R1| * |R2| * |R3| * RF(p1) * RF(p2) * RF(p3)$
c. $RF(p1) * RF(p2) * RF(p3) - (|R1| + |R2| + |R3|)$
d. $|R1| + |R2| + |R3| + RF(p1) + RF(p2) + RF(p3)$
e. none of the above answers is correct.

15. Encode the data *de gustibus non disputandum* using the secret encryption key *metallica* and the table of codes:

	a	b	c	d	e	f	G	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	-
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

(0.2p / question)

1p of