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

I. Choose the correct answer(s) for multiple choice questions 1 to 11. Each question has at least one correct answer. For questions 12 and 13, write only the final result(s). **Enter all answers in file Answers.docx. Only the answers in this file will be considered for the grade.**

1. Consider schedule S below:

T4	**T5**	**T6**
		read(C)
		C = C + 300
		write(C)
	read(D)	
	D = D + 20	
	write(D)	
read(C)		
C = C + 10		
write(C)		
		read(D)
		D = D - 100
		write(D)
	read(C)	
	C = C + 10	
	write(C)	
read(D)		
D = D + 20		
write(D)		

time

Before the execution above, C = 300 and D = 400.

a. S is serializable. A serial schedule that's equivalent to S (in its effect on the database) is: T5 followed by T6 followed by T4.
b. S is not serializable.
c. S is serializable. A serial schedule that's equivalent to S (in its effect on the database) is: T6 followed by T4 followed by T5.
d. S is serializable. A serial schedule that's equivalent to S (in its effect on the database) is: T6 followed by T5 followed by T4.
e. None of the above answers is correct.

2. Let R be a relation with 2000 pages. There are 50 pages in the buffer pool. R is sorted with *external merge sort*. Then:
a. Pass 0 produces 2 runs.
b. Each run produced in pass 0 has 2000 pages.
c. Pass 0 produces 10 runs.
d. Each run produced in pass 0 has 50 pages.
e. None of the above answers is correct.

3. In the Two-Phase Locking protocol:
a. A transaction can write object O without acquiring an X lock on O.
b. A transaction can't release locks before it completes execution.
c. A transaction can read object O without acquiring an S lock on O.
d. Once a transaction releases a lock, it cannot request other locks.
e. None of the above answers is correct.

4. In SQL Server:
a. Under READ UNCOMMITTED, exclusive locks are released at the end of the transaction.
b. Under READ UNCOMMITTED, shared locks are released at the end of the transaction.
c. Dirty reads can't occur under SERIALIZABLE.
d. Unrepeatable reads can occur under READ COMMITTED.
e. None of the above answers is correct.

5. Let T be a *linear* tree. Then for each join node N in T:
a. The left child of N must be a base relation.
b. The third child of N must be a base relation.
c. All children of N must be base relations.
d. The right child of N must be a base relation.
e. None of the above answers is correct.

6. A query block:
a. Has at most one GROUP BY clause.
b. Has exactly two HAVING clauses.
c. Has exactly one FROM clause.
d. Has at most one WHERE clause.
e. None of the above answers is correct.

7. I is an index with search key <A, B, C>. If I is a:
a. B+ tree index, I matches condition A=3.
b. hash index, I matches condition A=1 AND B=2 AND D=3.
c. hash index, I matches condition A=1 AND B>2 AND C=3.
d. B+ tree index, I matches condition A > 10.
e. None of the above answers is correct.

8. The Hash Join algorithm is an instance of the:
a. iteration technique
b. indexing technique
c. derivation technique
d. partitioning technique
e. None of the above answers is correct.

9. In vertical fragmentation:
a. Fragmentation is performed with the *selection* operator.
b. Fragmentation is performed with the *projection* operator.
c. The reconstruction operator is *union*.
d. The reconstruction operator is *intersection*.
e. None of the above answers is correct.

10. Consider schedule S below (all transactions commit):

T1	**T2**	**T3**
read(A)		
	read(B)	
read(C)		
	write(A)	
write(B)		
		read(C)
		write(D)
write(C)		
		read(A)

time

a. S is conflict serializable.
b. (write(T2, A), read(T3, A)) belongs to the conflict relation of S.
c. (read(T1, C), read(T3, C)) belongs to the conflict relation of S.
d. the following serial schedule is conflict equivalent with S: T1 followed by T2 followed by T3.
e. None of the above answers is correct.

11. $S \otimes_{S.PID=P.PID} P$ is evaluated with *Block Nested Loops Join.* S is the outer relation. There are 202 pages in the buffer pool. S has 700 pages. Then P is scanned:
a. 5 times
b. 1 time
c. 4 times
d. 10 times
e. None of the above answers is correct.

12. Encode the data *gutta cavat lapidem non vi, sed saepe cadendo* using the secret encryption key *anaaslan* and the table of codes below. Let *M* be the obtained string. Write the substring of *M* that consists of characters on positions <17, 18, 19> (the first character in *M* is on position 1).

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

13. Let T1 and T2 be 2 relations. T1 has 90.000 records; a page can hold 100 T1 records. T2 has 300.000 records; a page can hold 50 T2 records.
a. Compute the cost of $T2 \otimes_{T2.ID=T1.IDT2} T1$ using *page-oriented nested loops join*. T2 is the outer relation.
b. 100 buffer pages are available. Compute the cost of $T2 \otimes_{T2.ID=T1.IDT2} T1$ using *sort-merge join*. T1 and T2 are not sorted beforehand. T1 is the outer relation. Use *external merge sort* to sort T1 and T2. Assume each partition is scanned once during the merging phase of *sort-merge join.*
c. 300 buffer pages are available. T1 is not sorted beforehand. Compute the cost of $\pi_{ID,C1,C2}(T1)$ using *projection based on sorting* (basic version, without improvement). Use *external merge sort.* The size of a tuple in the result of $\pi_{ID,C1,C2}(T1)$ is 1/5 times the size of a tuple in T1.

II. Think of an application that's powered by a relational database. In this context:

a. Draw the database diagram (at least 4 interrelated tables, with primary keys and foreign keys).

b. Describe one real-world scenario in which different users are trying to access and manipulate the data concurrently, along with the problems that can occur.

c. Write a query with one SELECT clause, one FROM clause, one WHERE clause and one JOIN clause. The SELECT clause must contain two columns. The WHERE clause must have two terms of the form *AttributeName Operator Value*, where *Operator* $\in \{<, \leq, =, >, \geq, <>\}$; the terms must be connected by AND. Draw an evaluation plan for the query and describe the query's evaluation (with concrete algorithms and costs).