

EKSAMENSFORSIDE

Skoleeksamen med tilsyn

Emnekode: DAT2000R	Emnenavn: Databases 2	
Dato: 22.11.2021	Tid fra / til: 09:00—13:00	Ant. timer: 4 timer
Emneansvarlig: Ali Chelli		
Campus: Ringerike	Fakultet: School of Business	
Antall oppgaver: 6	Antall vedlegg: 0	Ant. sider inkl. forside og vedlegg: 7 pages including the cover page
Tillatte hjelpemidler (jfr. emneplan): Ingen		
Opplysninger om vedlegg: Ingen		
Merknader:		

Ved eksamen på papir:

Kryss av for type eksamenspapir

Ruter ☐Linjer ☒**KANDIDATEN MÅ SELV KONTROLLERE AT OPPGAVESETTET ER FULLSTENDIG**

Problem 1: Choose the correct answer for the following sentences (20 points)

- 1) _____ means that the effects of a committed transaction are permanent and must not be lost because of later failure.
 - a. Atomicity
 - b. Consistency
 - c. Isolation
 - d. Durability
- 2) In _____ the tags are defined by the user.
 - a. JSON
 - b. XML
 - c. HTML
 - d. Log files
- 3) _____ are applied through GRANT and REVOKE commands.
 - a. Symmetric encryption algorithms
 - b. Mandatory Access Control (MAC)
 - c. Asymmetric encryption algorithms
 - d. Discretionary Access Controls (DAC)
- 4) In _____: Entities, such as users, are represented by nodes, whereas the connections between entities dictate how they are related.
 - a. Key-Value Databases
 - b. Wide-column Databases
 - c. Graph Databases
 - d. Document Databases
- 5) _____ provides scalable and reliable storage and allows to add commodity hardware to increase your storage capacity.
 - a. JSON
 - b. MapReduce
 - c. HDFS
 - d. YARN
- 6) _____ is the process of restoring the database to a consistent state in the event of failure.
 - a. Access control
 - b. Transaction management
 - c. Database recovery
 - d. Two Phase locking
- 7) In _____ records appear randomly distributed across the available file space.
 - a. Sequential files
 - b. B+-trees
 - c. Heap files
 - d. Hash files

- 8) XML and JSON databases are _____.
a. Key-Value Databases
b. Document Databases
c. Graph Databases
d. Wide-column Databases
- 9) _____ means that an object contains both the data structure and the set of methods that can be used to manipulate it.
a. Class
b. Inheritance
c. Encapsulation
d. Polymorphism.
- 10) _____ is a situation where a successfully completed update is overridden by another transaction.
a. Deadlock
b. Lost update
c. Inconsistent Analysis
d. Uncommitted dependency

Problem 2: Physical Storage (17 points)

1. Discuss the differences between heap and sequential file organizations by indicating:
 - a. Which file organization method is faster when inserting a record? Explain the reason behind that.
 - b. Which method is faster when searching for a record? Explain the reason behind that.
2. Discuss the differences between primary and secondary indexes.
3. Question on B+-tree:
 - a. We consider a B+-tree of order $M=3$. For this B+-tree, what is the maximum number of keys per node? and what is the maximum number of children per node?
 - b. Build a B+-tree of order $M=3$; insert the elements 7, 12, 15; and draw the resulting tree. We call the resulting tree "Tree 2".
 - c. Insert the value 16 in Tree 2 and draw the resulting tree. We call the obtained tree "Tree 3".
 - d. Insert the value 25 in Tree 3 and draw the resulting tree.

Problem 3: Triggers (10 points)

We Consider the tables Customers and Orders with the following structures:

Customers (customer_id, name, last_order_id)

Orders (order_id, item_id, quantity, price, customer_id)

The following SQL code is used to create the tables Customers and Orders:

```
CREATE TABLE Customers
(
    customer_id integer NOT Null,
    name character varying (20) NOT Null,
    last_order_id integer,
    PRIMARY KEY (customer_id)
)
```

```
CREATE TABLE Orders
(
    order_id integer NOT Null,
    item_id integer,
    quantity integer,
    price integer,
    customer_id integer,
    PRIMARY KEY (order_id)
)
```

Now we insert into the Customers table three rows using the following SQL code:

```
INSERT INTO Customers (customer_id, name) VALUES (1, 'Bob');
```

```
INSERT INTO Customers (customer_id, name) VALUES (2, 'Sally');
```

```
INSERT INTO Customers (customer_id, name) VALUES (3, 'Fred');
```

We now use the following SQL statement to display the content of the table Customers:

```
SELECT * FROM Customers;
```

1. Provide the values for the fields customer_id, name, and last_order_id for row 1, row 2, and row 3.
2. Create a trigger called newOrder that fills the missing value last_order_id in the table Customers based on the insertions that are made in the Orders table. Write the SQL code for the trigger newOrder.

Now we insert into the Orders table three records using the following SQL code:

```
INSERT INTO Orders (order_id, item_id, quantity, price, customer_id)
VALUES (1, 102, 5, 500, 3);
INSERT INTO Orders (order_id, item_id, quantity, price, customer_id)
VALUES (2, 103, 20, 50, 1);
INSERT INTO Orders (order_id, item_id, quantity, price, customer_id)
VALUES (3, 104, 15, 100, 2);
```

After that, we display the content of the table Customers using the following SQL statement:

```
SELECT * FROM Customers
```

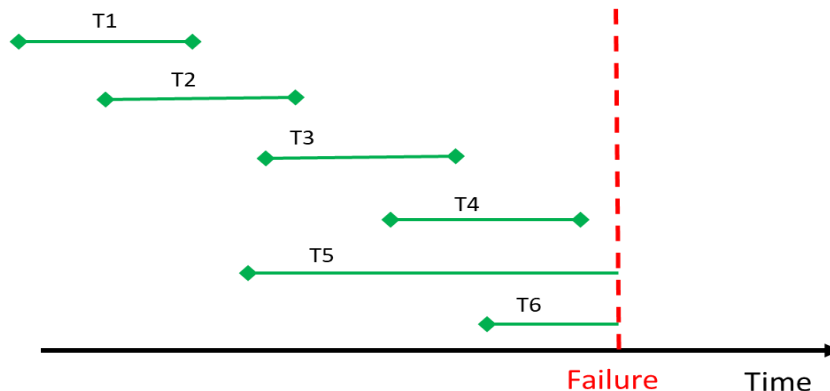
3. For each of the three rows in the Customers table, provide the values for the fields:
- customer_id
 - name
 - last_order_id

Problem 4: Transactions Management (18 points)

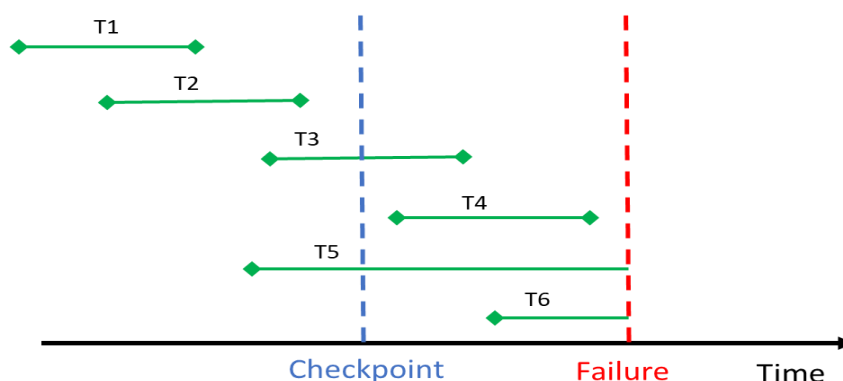
- Provide the definition of the ACID properties.
- We consider the following notation: W1(A) is a write operation performed by Transaction 1 on data item A. R2(B) is a read operation performed by Transaction 2 on data item B. For the following list of operations, indicate for each case if we have a conflict or not, and justify your answer.
 - R1(A), R2(A)
 - R1(B), W2(B)
 - W1(A), W2(B)
 - W1(A), R2(B)
- For each of the following schedules, indicate if it is conflict serializable or not, and justify your answer using a precedence graph.
 - S1: R1(A), W1(B), R2(A), W2(A), R1(B), R2(B)
 - S2: W1(B), R1(A), W2(A), R2(B), R1(B), W1(B)
 - S3: W2(A), R2(B), R1(B), W2(B), R1(A), W3(B)

Problem 5: Backup and Recovery (20 points)

1. Explain the difference between volatile storage and non-volatile storage. (Hint: compare the behavior of volatile storage and non-volatile storage after a power outage).
2. Is a log file saved on volatile storage or non-volatile storage? Explain the reason for that.
3. Consider the scenario shown in **Figure 1** where a system failure occurs as indicated by the vertical red dashed line. After the system restart, determine which transactions must be undone and which transaction must be redone.

**Figure 1**

4. Consider now the scenario shown in **Figure 2** where a checkpoint is indicated by the vertical blue dashed line. After the system restart, determine which transactions must be undone and which transaction must be redone. Moreover, specify the transactions for which we must do nothing.
5. Explain what a checkpoint is?
6. Based on your answers to questions 3 and 4, conclude the benefit of using a checkpoint.

**Figure 2**

Problem 6: XML, JSON, Big Data, and NoSQL (15 points)

1. What are the five Vs of Big Data?
2. Mention the major four types of NoSQL databases.
3. Mention one advantage of JSON compared to XML.
4. For each of the following JSON examples, indicate whether it has a valid syntax or not. For the examples with invalid syntax, indicate the syntax error.
 - a. `{"title":"Star Wars", "quotes":["Use The Force","These are not the Droids you are looking for"],"director":"George Lucas"}`
 - b. `{"name":"Fred Flintstone";"job":"Doctor";"wife":"Kristine"}`
 - c. `{item:"paper", Qty:120, status: D, tags: ["Red", "Blank"]}`
 - d. `{"city","New York", "population", 7999034}`