

EXAMINATION PAPER

Exam in: INF-2700 Database Systems

Date: Thursday 03.12.2015

Time: Kl 15:00 - 19:00

Place: Åsgårdveien 9

Approved aids: None

The exam contains 5 pages, including this cover page.

Contact person: Weihai Yu.

Phone: 41429077

NB! It is not allowed to submit scratch paper along with the answer sheets

Question 1 (40%)

Below are some database tables with example data for a shopping application.

- Products

pid	pname	price
p01	cup	49
p02	hat	99
p03	pen	19

- Customers

cid	cname	city
c01	ida	oslo
c02	ida	alta
c03	ole	alta
c04	tom	oslo

- Orders

oid	cid	pid	quantity
101	c01	p01	3
102	c01	p02	5
103	c01	p02	1
111	c02	p01	2
112	c02	p02	2
113	c02	p03	2
121	c03	p01	1
122	c03	p01	9
131	c04	p01	1

The *primary keys* of the tables are in **bold** text.

Foreign keys in Orders:

- cid: references cid of Customers
- pid: references pid of Products

Write queries to find the required information.

Queries 1–5 must be formulated in *both relational algebra and SQL*.

Queries 6–10 need only be formulated in *SQL*.

Note: In the result tables of your SQL queries, there should be *no* identical (duplicate) rows.

Relational algebra *and* SQL (1–5):

1. Names of all customers.

The result for the example database is:

cname
ida
ole
tom

2. Products with price lower than 50 Kr.

The result for the example database is:

pid	pname	price
p01	cup	49
p03	pen	19

3. Orders from Ida in Alta.

The result for the example database is:

oid	pname	quantity
111	cup	2
112	hat	2
113	pen	2

4. Cids of customers who ordered both p01 and p02.

The result for the example database is:

cid
c01
c02

5. Cids of customers who only ordered one kind of product.

The result for the example database is:

cid
c03
c04

SQL *only* (6–10):

6. Number of different kinds of products.

The result for the example database is:

numberOfProducts
3

7. Total price of orders from customers in Oslo.

The result for the example database is:

totalPriceFromOslo
790

8. Cids of customers who ordered more than two kinds of products.

The result for the example database is:

cid	numberOfOrderedProducts
c02	3

9. Cids of customers who ordered the most expensive product.

The result for the example database is:

cid
c01
c02

10. Cids of customers who ordered all products that c01 ordered.

The result for the example database is:

cid
c02

Question 2 (20%)

Now consider the physical data processing for the database in Question 1.

We decide to organize the database data as below:

- Table Products is organized with hash on pid.
- Table Customers is organized with hash on cid.
- Table Orders is organized with hash on cid. In addition, there is a hash index on pid.

Answer the following questions.

1. Sketch how Orders data are organized with a figure and some brief description.
2. We are going to make a natural join of the tables Customers and Orders $Customers \bowtie Orders$, with the *hash join* algorithm.
Describe how the algorithm works.
3. What is the primary performance overhead of database systems in general?
What is the performance overhead of $Customers \bowtie Orders$ with hash join? (You should make reasonable assumptions of data sizes.)

Question 3 (20%)

Answer the following questions. Please explain the relevant concepts while answering the questions.

1. What is *functional dependency* $X \rightarrow Y$ of a relation instance r ?
For the example instance of table Customers in Question 1, check if the following functional dependencies are satisfied.
 - a) $cid \rightarrow cname$
 - b) $cname \rightarrow city$
 - c) $\{cname, city\} \rightarrow cname$
 - d) $\{cname, city\} \rightarrow cid$
2. What is *third normal form* (3NF)?
3. Given the relation schema $R(A, B, C)$, $F = \{A \rightarrow C, B \rightarrow C\}$. Explain why R is not in 3NF.
4. Can you decompose schema R into 3NF with the *3NF synthesis algorithm*?
5. Why do we need 3NF?
6. What problem may 3NF have?

Question 4 (20%)

1. What is an *ACID transaction*?
2. What is a *log* for transaction processing?
How is a log organized?
What are the operations on a log?
What are the performance costs of these operations?
3. What kinds of log records do you need to rollback individual transactions when the database system is up and running?
What information should the record contain?
4. What if you also want the database to recover from a system crash?

—END—