
King Fahd University of Petroleum and Minerals

Information and Computer Science Department

ICS 424: Advanced Database Systems
Second Semester (062)

FINAL EXAM (30%)

Switch off mobile phones and do not open until instructed to do so!!!

Write clearly, precisely, and briefly!!

ID:	
Name:	

Grades		
Section	Max	Scored
A	10	
B	15	
C	14	
D	19	
E	12	
TOTAL	70	

A. Questions from Chapters 19-21 (OODB and ORDB) (10)

1. Assume the following schema

Person

Student

Lecturer

Person attributes

- name (first name, last name)
- Date-of-birth

Student attributes

- major

Lecturer attributes

- advisees

- a. Write SQL statements which can create the tables of the above schema. (Use Oracle object-relational features.)

- b. Write one SQL statement which can insert into the `lecturer` table a row which belongs to a new lecturer who advises more than one student.

B. Questions from Chapter 25 (Distributed DBs) (15)

1. Explain how the concurrency control based on voting works. (5)

2. Assume your database is distributed between Dammam site and Jeddah site. Assume EMP table is in Dammam and DEPT is in Jeddah. The description of EMP and DEPT is as follows

EMP has 1,000,000 records and has the following attributes:

- ID 10 bytes
- fname 30 bytes
- lname 30 bytes
- Address 100 bytes
- Deptno 5 bytes
- Nationality 25 bytes

DEPT has 20 records and the following attributes

- DNO 5 bytes
- DName 15 bytes

Assume a person sitting in Jeddah wrote the following query

```
SELECT Fname  
FROM Dept, emp  
WHERE DeptNo = DNO
```

What is the best execution strategy of the above query, that results in the minimum communication cost? What is the communication cost of the strategy? **(10)**

C. Questions from Chapter 26 (XML) (14)

1. What is the difference between a valid XML document and a well formed XML document? (4)

2. Map the following DTD into relational schema. (10)

```
<!DOCTYPE books [  
  <!ELEMENT book (title, year, author+)>  
  <!ELEMENT author (fname, lname)>  
  <!ELEMENT fname (#PCDATA)>  
  <!ELEMENT lname (#PCDATA)>  
  <!ELEMENT title (#PCDATA)>  
  <!ELEMENT year (#PCDATA)>  
  <!ATTRIBUTE book ISBN CDATA #REQUIRED>  

```

D. Questions from Chapter 27 (Data mining) (19)

1. What is the difference between classification and clustering? (4)

2. Write the steps of basic clustering algorithm. (You can use pseudo code). (5)

3. Apply the Apriori algorithm to the following data set

(10)

Transaction	Items
T1	milk, bread
T2	milk, bread
T3	milk, juice
T4	bread, juice, coffee, eggs
T5	milk, eggs
T6	milk, bread, Juice

The set of items is (milk, bread, coffee, juice, eggs). Use **0.5** for the minimum support value. (Show all your steps)

E. Questions from Chapter 28 (Data warehousing) (12)

1. Give 2 reasons why we need to separate a data warehouse from a database. (4)

2. What is the use of a staging schema? (4)

3. Briefly explain the following 2 OLAP operations. (You can use an example.) (4)

a. Roll-up

b. Drill-down