

SWE3003 Introduction to Database Systems - Midterm Fall 2022

Student ID	Name

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Total
For Instructor/TA only,										

Academic Honor Pledge

I affirm that I will not at any time be involved with cheating or plagiarism while enrolled as a student Programming Language class at SungKyunKwan University.

I understand that violation of this code will result in penalties as severe as indefinite suspension from the university.

Your signature: _____

1. [20 pts] For each of the following statements, indicate whether it is TRUE or FALSE. You will get 2 points for each correct answer, -2 point for each incorrect answer, and 0 point for each answer left blank or both answers marked.

	T	F
(a) Given relations $R(A, B)$ and $S(A, C)$, the natural join $R \bowtie S$ is equal to $R \cap S$	<input type="checkbox"/>	✓
(b) In the WHERE clause of a SQL query, the condition 'John Doe' = NULL is evaluated to be true.	<input type="checkbox"/>	✓
(c) In the WHERE clause of a SQL query, the condition NULL = NULL is evaluated to be true.	<input type="checkbox"/>	✓
(d) It is possible to declare multiple different sets of attributes as UNIQUE in the same relation R.	✓	<input type="checkbox"/>
(e) If we define a foreign key in relation R, the DBMS checks the foreign-key constraint whenever a tuple in R is deleted.	✓	<input type="checkbox"/>
(f) A weak entity set does not have any attribute that belongs to its key. ...	<input type="checkbox"/>	✓
(g) If a functional dependency $A \rightarrow B$ holds in relation $R(A, B, C)$, then $AC \rightarrow B$ also holds.	✓	<input type="checkbox"/>
(h) If a relation R is in 3NF, R is also in BCNF.	<input type="checkbox"/>	✓
(i) The BCNF decomposition algorithm does not preserve functional dependencies of the initial relation R.	✓	<input type="checkbox"/>
(j) The leaf nodes of a B+tree are stored sequentially on disk.	<input type="checkbox"/>	✓

2. [10 pts] Consider the two relations:

- $R(A,B) = (p,1), (q,5), (r,3)$
- $S(B,C) = (1,p), (5,q), (5,r), (6,p)$

$R(A,B)$	$S(B,C)$
p 1	1 p
q 5	5 q
r 3	5 r
	6 p

Show the output relation for each of the following relational algebras.

- (a) $\Pi_{A,B}(R \bowtie S)$

answer:

p 1
q 5

.

- (b) $\sigma_{A=C}(R \bowtie S)$

answer:

p 1 p
q 5 q

.

3. [20 pts] Write each of the following queries in SQL for the given relations.

- Seller(seller_id, seller_name, city)
- Item(item_id, item_name, price)
- Buyer(buyer_id, buyer_name, street, city, phone)
- Order(order_no, seller_id, buyer_id, received, shipped)
- Order_details(order_no, item_id, qty)

a. Get order numbers for the orders where buyer and seller are in the same city.

answer (5 pts):

```
select order_number
from Seller join Order join Buyer
where Seller.city = Buyer.city

or

select order_number
from Seller natural join Order natural join Buyer
```

b. Get each seller's name and the total sales (i.e., total price of sold items) for each seller.

answer (5 pts):

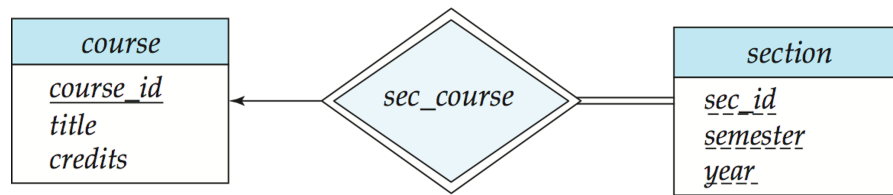
```
select seller_name, sum(price)
from Seller join Order join Order_details join Item
group by seller_id
```

c. Get the buyer IDs of buyers who have placed orders with ALL sellers from 'Suwon'.

answer (10 pts):

```
select distinct buyer_id
from Buyer as B
where not exists ( (select seller_id
                    from Seller
                    where city = 'Suwon')
                  except
                  (select seller_id
                   from Order O
                   where B.buyer_id = O.buyer_id));
```

4. [10 pts] Write SQL DDL statements for the following ER-diagram.
Note: all attributes are of VARCHAR(10) type.



answer:

```
CREATE TABLE course (  
    course_id VARCHAR(10) PRIMARY KEY,  
    title VARCHAR(10),  
    credits VARCHAR(10)  
);  
  
CREATE TABLE section (  
    course_id VARCHAR(10)  
        REFERENCES course(course_id) ON DELETE CASCADE, -- foreign key constraint  
    sec_id VARCHAR(10),  
    semester VARCHAR(10),  
    year VARCHAR(10),  
    PRIMARY KEY (course_id, sec_id, semester, year) -- this primary key  
        -- constraint associates weak entities with strong entities.  
);
```

5. [15 pts] Consider the following relations and functional dependencies (FDs) below.

- Prescription(doctor_id, doctor_name, patient_id, diagnosis)
- FD1: doctor_id \rightarrow doctor_name
- FD2: doctor_id, patient_id \rightarrow diagnosis
- FD3: patient_id, diagnosis \rightarrow doctor_id

(a) Decompose this relation into BCNF

answer:

A (doctor_id, doctor_name, patient_id)

B (doctor_id, patient_id, diagnosis)

FD2: 'doctor_id, patient_id \rightarrow diagnosis' violates BCNF,
so we need to decompose the schema into

A (doctor_id, doctor_name, patient_id)

FD1: doctor_id \rightarrow doctor_name

B (doctor_id, patient_id, diagnosis)

FD2: doctor_id, patient_id \rightarrow diagnosis

FD3: patient_id, diagnosis \rightarrow doctor_id

(doctor_id, patient_id) is a candidate key, i.e., super key.

(patient_id, diagnosis) is also a candidate key, i.e., super key.

So, FD2 and FD3 satisfy BCNF.

(b) Decompose the relation into 3NF.

answer:

Same...


(c) What are the advantage and disadvantage of 3NF compared to BCNF?

answer:

- BCNF does not always preserve dependency. Too much decomposition results in lossy join.
- 3NF guarantees lossless-join while preserving dependency, but it allows some redundancy and lots of NULL values.

6. [10 pts] Below is the state of blocks used for a heap file.

B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15
----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----

 used block

 free block

(a) Draw a free-space map and its second-level free-space map for the given heap file (B0~B15). Both the first-level and second-level free-space maps use 2 bits per entry.

answer:

2 bits: 0 - 3

1st level: the question is not clear as to how many entries are grouped.
so.. ignore the 1st level.

2nd level: 2 2 1 2 0 0

(b) Suppose block B1 is used to store records in the heap file. Show how the free-space maps will change.

answer:

2nd level: 1 2 1 2 0 0

7. [10 pts] Consider the following two tables - r and s are stored in a multitable clustering file.

A	B	C		B	D
X	1	a		1	a
X	1	b		2	a
Y	2	b		2	b
Y	4	c			
r				s	

- (a) Draw how the records in r and s will be laid out in the file.

answer:

X 1 a
X 1 b
1 a
Y 2 b
2 a
2 b
Y 4 c

OR

1 a
X 1 a
X 1 b

2 a
2 b
Y 2 b

Y 4 c

If 1s and 2s are grouped, consider the answer correct.

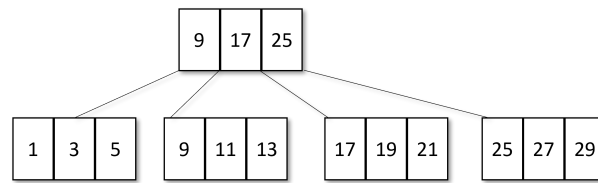
- (b) What is the main advantage and disadvantage of using multitable clustering file?

advantage:

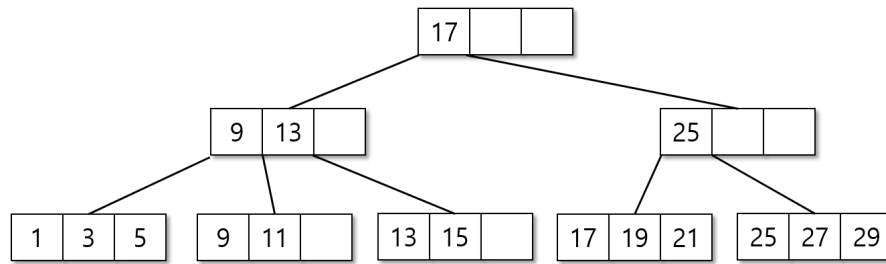
Good for join operations

disadvantage:
Bad for queries that access only one table.

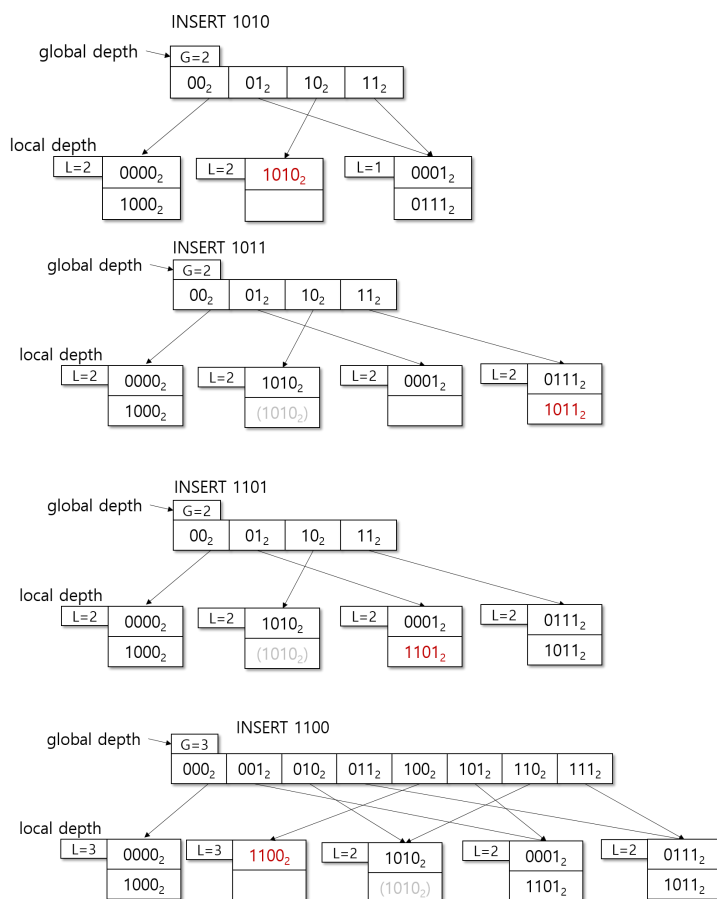
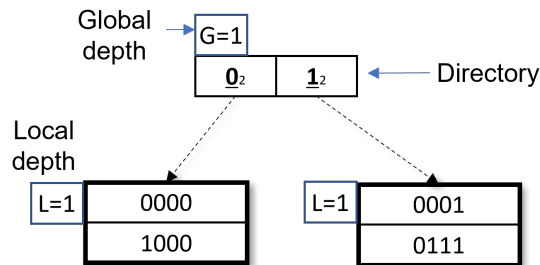
8. [10 pts] Below is the state of B+tree with degree 4. Draw the state of the B+tree after inserting 15.



answer:



9. [10 pts] Below is the state of extendible hash table after inserting four data entries. We use the least significant bits of keys and $\%2^G$ as the hash function. Draw a diagram showing the state of the extendible hash table after inserting the following 6 more data entries.
1010, 1011, 1010, 1101, 1100, 1001



answer:

