3 Databases (229)

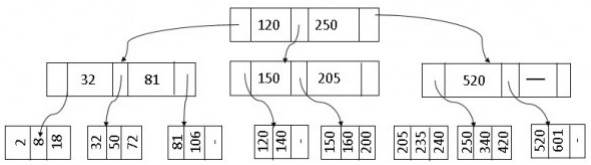
ER-model. Relational model:Relational algebra, Tuple calculus, SQL. Integrity constraints, Normal forms. File organization, Indexing (e.g., B and B+ trees). Transactions and concurrency control.

3.1 B Tree (26)

3.1.1 B Tree: GATE1989-12a

https://gateoverflow.in/91199

Fig.7 shows a $B+$ tree where only key values are indicated in the records. Each block can hold upto three records. A record with a key value 34 is inserted into the $B+$ tree. Obtain the modified $B+$ tree after insertion.



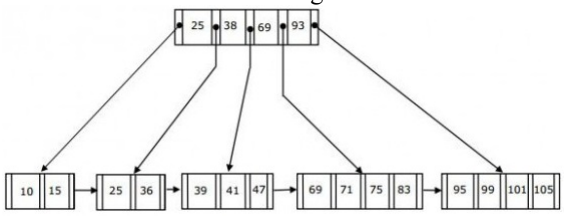
descriptive gate1989 databases b-tree

3.1.2 B Tree: GATE1994-14

https://gateoverflow.in/2510

Consider B^+ - tree of order d shown in figure. (A B^+ - tree of order d contains between d and $2d$ keys in each node)

A. Draw the resulting B^+ - tree after 100 is inserted in the figure below.



B. For a B^+ - tree of order d with n leaf nodes, the number of nodes accessed during a search is $O(-)$.

gate1994 databases b-tree normal

3.1.3 B Tree: GATE1997-19

https://gateoverflow.in/2279

A B^+ - tree of order d is a tree in which each internal node has between d and $2d$ key values. An internal node with M key values has $M+1$ children. The root (if it is an internal node) has between 1 and $2d$ key values. The distance of a node from the root is the length of the path from the root to the node. All leaves are at the same distance from the root. The height of the tree is the distance of a leaf from the root.

A. What is the total number of key values in the internal nodes of a B^+-tree with l leaves ($l \geq 2$)?
B. What is the maximum number of internal nodes in a B^+ - tree of order 4 with 52 leaves?
C. What is the minimum number of leaves in a B^+-tree of order d and height $h (h \geq 1)$?

gate1997 databases b-tree normal

3.1.4 B Tree: GATE1999-1.25

https://gateoverflow.in/1478

Which of the following is correct?

A. B-trees are for storing data on disk and B^+ trees are for main memory.
B. Range queries are faster on B^+ trees.
C. B-trees are for primary indexes and B^+ trees are for secondary indexes.
D. The height of a B^+ tree is independent of the number of records.

gate1999 databases b-tree normal

3.1.5 B Tree: GATE1999-21

https://gateoverflow.in/1520

Consider a B-tree with degree m, that is, the number of children, c, of any internal note (except the root) is such that $m \leq c \leq 2m-1$. Derive the maximum and minimum number of records in the leaf nodes for such a B-tree with height $h, h \geq 1$. (Assume that the root of a tree is at height 0).

gate1999 databases b-tree normal

3.1.6 B Tree: GATE2000-1.22, UGCNET-June2012-II-11

https://gateoverflow.in/646

B^+-trees are preferred to binary trees in databases because

A. Disk capacities are greater than memory capacities
B. Disk access is much slower than memory access
C. Disk data transfer rates are much less than memory data transfer rates
D. Disks are more reliable than memory

gate2000 databases b-tree normal ugcnetjune2012ii

3.1.7 B Tree: GATE2000-21

https://gateoverflow.in/692

(a) Suppose you are given an empty B+- tree where each node (leaf and internal) can store up to 5 key values. Suppose values 1, 2,.....10 are inserted, in order, into the tree. Show the tree pictorially

i. after 6 insertions, and
ii. after all 10 insertions

Do NOT show intermediate stages.

(b) Suppose instead of splitting a node when it is full, we try to move a value to the left sibling. If there is no left sibling, or the left sibling is full, we split the node. Show the tree after values 1, 2,....., 9 have been inserted. Assume, as in (a) that each node can hold up to 5 keys.

(c) In general, suppose a B+- tree node can hold a maximum of m keys,and you insert a long sequence of keys in increasing order. Then what approximately is the average number of keys in each leaf level node.

i. in the normal case, and
ii. with the insertion as in (b).

gate2000 databases b-tree normal descriptive

3.1.8 B Tree: GATE2002-17

https://gateoverflow.in/870

a. The following table refers to search items for a key in B-trees and B^+ trees.

B-tree		**B^+-tree**	
Successful search	Unsuccessful search	Successful search	Unsuccessful search
X_1	X_2	X_3	X_4

A successful search means that the key exists in the database and unsuccessful means that it is not present in the database. Each of the entries X_1, X_2, X_3 and X_4 can have a value of either Constant or Variable. Constant means that the search time is the same, independent of the specific key value, where variable means that it is dependent on the specific key value chosen for the search.
Give the correct values for the entries X_1, X_2, X_3 and X_4 (for example $X_1 = \text{Constant}$, $X_2 = \text{Constant}$, $X_3 = \text{Constant}, X_4 = \text{Constant}$)

b. Relation $R(A, B)$ has the following view defined on it:

```
CREATE VIEW V AS
(SELECT R1.A,R2.B
FROM R AS R1, R as R2
WHERE R1.B=R2.A)
```

i. The current contents of relation R are shown below. What are the contents of the view V?

A	B
1	2
2	3
2	4
4	5
6	7
6	8
9	10

ii. The tuples $(2, 11)$ and $(11, 6)$ are now inserted into R. What are the *additional* tuples that are inserted in V?

gate2002 databases b-tree normal descriptive

3.1.9 B Tree: GATE2002-2.23, UGCNET-June2012-II-26

https://gateoverflow.in/853

A B^+ - tree index is to be built on the *Name* attribute of the relation *STUDENT*. Assume that all the student names are of length 8 bytes, disk blocks are of size 512 bytes, and index pointers are of size 4 bytes. Given the scenario, what would be the best choice of the degree (i.e. number of pointers per node) of the B^+ - tree?

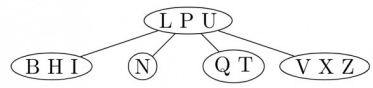
A. 16 B. 42 C. 43 D. 44

gate2002 databases b-tree normal ugcnetjune2012ii

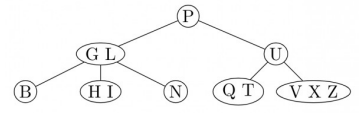
3.1.10 B Tree: GATE2003-65

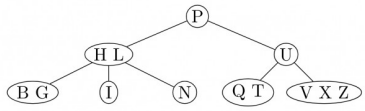
https://gateoverflow.in/952

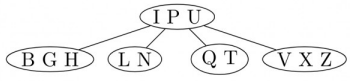
Consider the following $2-3-4$ tree (i.e., B-tree with a minimum degree of two) in which each data item is a letter. The usual alphabetical ordering of letters is used in constructing the tree.



What is the result of inserting G in the above tree?

A.


B.


C.


D. None of the above

gate2003 databases b-tree normal

3.1.11 B Tree: GATE2004-52

https://gateoverflow.in/1048

The order of an internal node in a $B+$ tree index is the maximum number of children it can have. Suppose that a child pointer takes 6 bytes, the search field value takes 14 bytes, and the block size is 512 bytes. What is the order of the internal node?

A. 24 B. 25 C. 26 D. 27

gate2004 databases b-tree normal

3.1.12 B Tree: GATE2004-IT-79

https://gateoverflow.in/3723

Consider a table T in a relational database with a key field K. A B-tree of order p is used as an access structure on K, where p denotes the maximum number of tree pointers in a B-tree index node. Assume that K is 10 $bytes$ long; disk block size is 512 $bytes$; each data pointer P_D is 8 bytes long and each block pointer P_B is 5 bytes long. In order for each B-tree node to fit in a single disk block, the maximum value of p is

A. 20 B. 22 C. 23 D. 32

gate2004-it databases b-tree normal

3.1.13 B Tree: GATE2005-28

https://gateoverflow.in/1364

Which of the following is a key factor for preferring B^+-trees to binary search trees for indexing database relations?

A. Database relations have a large number of records
B. Database relations are sorted on the primary key
C. B^+-trees require less memory than binary search trees
D. Data transfer form disks is in blocks

gate2005 databases b-tree normal

3.1.14 B Tree: GATE2005-IT-23, ISRO2017-67

https://gateoverflow.in/3768

A B-Tree used as an index for a large database table has four levels including the root node. If a new key is inserted in this index, then the maximum number of nodes that could be newly created in the process are

A. 5 B. 4 C. 3 D. 2

gate2005-it databases b-tree normal isro2017

3.1.15 B Tree: GATE2006-IT-61

https://gateoverflow.in/3605

In a database file structure, the search key field is $9\ bytes$ long, the block size is $512\ bytes$, a record pointer is $7\ bytes$ and a block pointer is $6\ bytes$. The largest possible order of a non-leaf node in a $B+$ tree implementing this file structure is

A. 23 B. 24 C. 34 D. 44

gate2006-it databases b-tree normal

3.1.16 B Tree: GATE2007-63, ISRO2016-59

https://gateoverflow.in/1261

The order of a leaf node in a B $^+$ - tree is the maximum number of (value, data record pointer) pairs it can hold. Given that the block size is $1K\ bytes$, data record pointer is $7\ bytes$ long, the value field is $9\ bytes$ long and a block pointer is $6\ bytes$ long, what is the order of the leaf node?

A. 63 B. 64 C. 67 D. 68

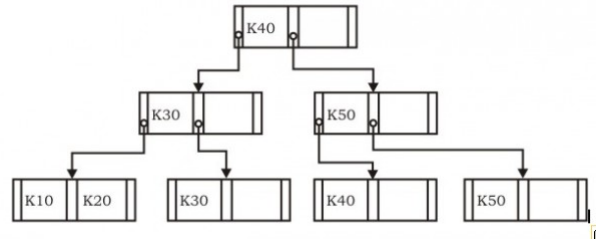
leaf node has different formula
n*(K+RP)+BP <= Block size

gate2007 databases b-tree normal isro2016

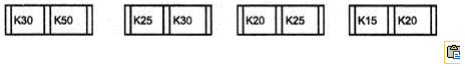
3.1.17 B Tree: GATE2007-IT-84

https://gateoverflow.in/3536

Consider the B^+ tree in the adjoining figure, where each node has at most two keys and three links.



Keys $K15$ and then $K25$ are inserted into this tree in that order. Exactly how many of the following nodes (disregarding the links) will be present in the tree after the two insertions?



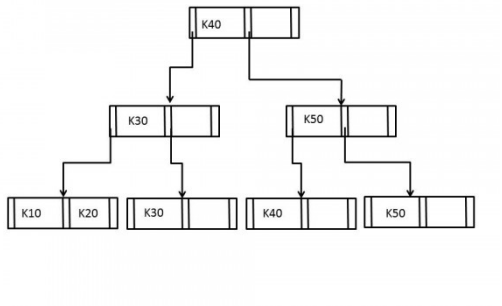
A. 1 B. 2 C. 3 D. 4

gate2007-it databases b-tree normal

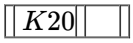
3.1.18 B Tree: GATE2007-IT-85

https://gateoverflow.in/3537

Consider the B^+ tree in the adjoining figure, where each node has at most two keys and three links.



Now the key $K50$ is deleted from the B^+ tree resulting after the two insertions made earlier. Consider the following statements about the B^+ tree resulting after this deletion.

i. The height of the tree remains the same.
ii. The node


(disregarding the links) is present in the tree.
iii. The root node remains unchanged (disregarding the links) 1

Which one of the following options is true ?

A. Statements (i) and (ii) are true
B. Statements (ii) and (iii) are true
C. Statements (iii) and (i) are true
D. All the statements are false

gate2007-it databases b-tree normal

3.1.19 B Tree: GATE2008-41

https://gateoverflow.in/453

A B-tree of order 4 is built from scratch by 10 successive insertions. What is the maximum number of node splitting operations that may take place?

A. 3 B. 4 C. 5 D. 6

gate2008 databases b-tree normal

3.1.20 B Tree: GATE2009-44

https://gateoverflow.in/1330

The following key values are inserted into a $B+$ - tree in which order of the internal nodes is 3, and that of the leaf nodes is 2, in the sequence given below. The order of internal nodes is the maximum number of tree pointers in each node, and the order of leaf nodes is the maximum number of data items that can be stored in it. The $B+$ - tree is initially empty

$10, 3, 6, 8, 4, 2, 1$

The maximum number of times leaf nodes would get split up as a result of these insertions is

A. 2 B. 3 C. 4 D. 5

gate2009 databases b-tree normal

3.1.21 B Tree: GATE2010-18

https://gateoverflow.in/2191

Consider a B^+-tree in which the maximum number of keys in a node is 5. What is the minimum number of keys in any non-root node?

A. 1 B. 2 C. 3 D. 4

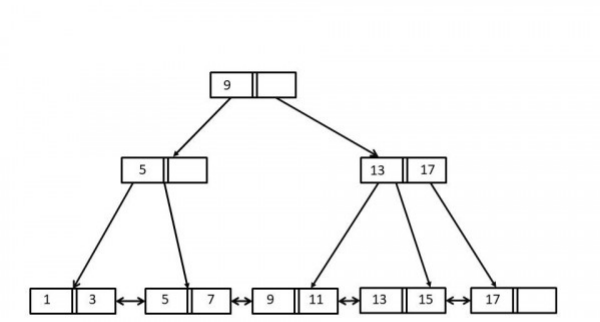
gate2010 databases b-tree easy

3.1.22 B Tree: GATE2015-2-6

https://gateoverflow.in/8052

With reference to the B+ tree index of order 1 shown below, the minimum number of nodes (including the Root node) that must be fetched in order to satisfy the following query. "Get all records with a search key greater than or equal to 7

and less than 15 " is _____.



gate2015-2 databases b-tree normal numerical-answers

3.1.23 B Tree: GATE2015-3-46

https://gateoverflow.in/8555

Consider a B+ tree in which the search key is 12 byte long, block size is 1024 byte, recorder pointer is 10 byte long and the block pointer is 8 byte long. The maximum number of keys that can be accommodated in each non-leaf node of the tree is _____.

gate2015-3 databases b-tree normal numerical-answers

3.1.24 B Tree: GATE2016-2-21

https://gateoverflow.in/39569

B+ Trees are considered BALANCED because.

A. The lengths of the paths from the root to all leaf nodes are all equal.
B. The lengths of the paths from the root to all leaf nodes differ from each other by at most 1.
C. The number of children of any two non-leaf sibling nodes differ by at most 1.
D. The number of records in any two leaf nodes differ by at most 1.

gate2016-2 databases b-tree normal

3.1.25 B Tree: GATE2017-2-49

https://gateoverflow.in/118561

In a B^+ Tree , if the search-key value is 8 bytes long , the block size is 512 bytes and the pointer size is 2 B , then the maximum order of the B^+ Tree is ____

gate2017-2 databases b-tree numerical-answers normal

3.1.26 B Tree: GATE2019-14

https://gateoverflow.in/302834

Which one of the following statements is NOT correct about the B+ tree data structure used for creating an index of a relational database table?

A. B+ Tree is a height-balanced tree
B. Non-leaf nodes have pointers to data records
C. Key values in each node are kept in sorted order
D. Each leaf node has a pointer to the next leaf node

gate2019 databases b-tree

3.2 Candidate Keys (5)

3.2.1 Candidate Keys: GATE1994-3.7

https://gateoverflow.in/2493

An instance of a relational scheme $R(A, B, C)$ has distinct values for attribute A. Can you conclude that A is a candidate key for R?

gate1994 databases easy candidate-keys

3.2.2 Candidate Keys: GATE2011-12

https://gateoverflow.in/2114

Consider a relational table with a single record for each registered student with the following attributes:

1. Registration_Num: Unique registration number for each registered student
2. UID: Unique identity number, unique at the national level for each citizen
3. BankAccount_Num: Unique account number at the bank. A student can have multiple accounts or joint accounts. This attribute stores the primary account number.
4. Name: Name of the student
5. Hostel_Room: Room number of the hostel

Which of the following options is **INCORRECT**?

A. BankAccount_Num is a candidate key
B. Registration_Num can be a primary key
C. UID is a candidate key if all students are from the same country
D. If S is a super key such that $S \cap$ UID is NULL then $S \cup$ UID is also a superkey

gate2011 databases normal candidate-keys

3.2.3 Candidate Keys: GATE2014-2-21

https://gateoverflow.in/1978

The maximum number of superkeys for the relation schema $R(E, F, G, H)$ with E as the key is _____.

gate2014-2 databases numerical-answers easy candidate-keys

3.2.4 Candidate Keys: GATE2014-2-22

https://gateoverflow.in/1980

Given an instance of the STUDENTS relation as shown as below

StudentID	StudentName	StudentEmail	StudentAge	CPI
2345	Shankar	shankar@math	X	9.4
1287	Swati	swati@ee	19	9.5
7853	Shankar	shankar@cse	19	9.4
9876	Swati	swati@mech	18	9.3
8765	Ganesh	ganesh@civil	19	8.7

For (StudentName, StudentAge) to be a key for this instance, the value X should NOT be equal to_____.

gate2014-2 databases numerical-answers easy candidate-keys

3.2.5 Candidate Keys: GATE2014-3-22

https://gateoverflow.in/2056

A *prime attribute* of a relation scheme R is an attribute that appears

A. in all candidate keys of R
B. in some candidate key of R
C. in a foreign key of R
D. only in the primary key of R

gate2014-3 databases easy candidate-keys

3.3 Conflict Serializable (1)

3.3.1 Conflict Serializable: GATE2017-2-44

https://gateoverflow.in/118640

Two transactions T_1 and T_2 are given as

$T_1 : r_1(X)w_1(X)r_1(Y)w_1(Y)$

$T_2 : r_2(Y)w_2(Y)r_2(Z)w_2(Z)$

where $r_i(V)$ denotes a $read$ operation by transaction T_i on a variable V and $w_i(V)$ denotes a $write$ operation by transaction T_i on a variable V. The total number of conflict serializable schedules that can be formed by T_1 and T_2 is _____

gate2017-2 databases transactions numerical-answers conflict-serializable

3.4 Data Independence (1)

3.4.1 Data Independence: GATE1994-3.11

https://gateoverflow.in/2497

State True or False with reason

Logical data independence is easier to achieve than physical data independence

gate1994 databases normal data-independence

3.5 Database Normalization (27)

3.5.1 Database Normalization: GATE1994-3.6

https://gateoverflow.in/2492

State True or False with reason

There is always a decomposition into Boyce-Codd normal form (BCNF) that is lossless and dependency preserving.

gate1994 databases database-normalization easy

3.5.2 Database Normalization: GATE1995-26

https://gateoverflow.in/2665

Consider the relation scheme $R(A, B, C)$ with the following functional dependencies:

- $A, B \rightarrow C$,
- $C \rightarrow A$

A. Show that the scheme R is in $3NF$ but not in BCNF.
B. Determine the minimal keys of relation R.

gate1995 databases database-normalization normal

3.5.3 Database Normalization: GATE1997-6.9

https://gateoverflow.in/2265

For a database relation $R(a, b, c, d)$, where the domains a, b, c, d include only atomic values, only the following functional dependencies and those that can be inferred from them hold

- $a \rightarrow c$
- $b \rightarrow d$

This relation is

A. in first normal form but not in second normal form
B. in second normal form but not in first normal form
C. in third normal form
D. none of the above

gate1997 databases database-normalization normal

3.5.4 Database Normalization: GATE1998-1.34

https://gateoverflow.in/1671

Which normal form is considered adequate for normal relational database design?

A. $2NF$ B. $5NF$ C. $4NF$ D. $3NF$

gate1998 databases database-normalization easy

3.5.5 Database Normalization: GATE1998-26

https://gateoverflow.in/1741

Consider the following database relations containing the attributes

- Book_id
- Subject_Category_of_book
- Name_of_Author
- Nationality_of_Author

With Book_id as the primary key.

a. What is the highest normal form satisfied by this relation?
b. Suppose the attributes Book_title and Author_address are added to the relation, and the primary key is changed to {Name_of_Author, Book_title}, what will be the highest normal form satisfied by the relation?

gate1998 databases database-normalization normal

3.5.6 Database Normalization: GATE1999-1.24

https://gateoverflow.in/1477

Let $R = (A, B, C, D, E, F)$ be a relation scheme with the following dependencies $C \rightarrow F, E \rightarrow A, EC \rightarrow D, A \rightarrow B$. Which one of the following is a key for R?

A. CD B. EC C. AE D. AC

gate1999 databases database-normalization easy

3.5.7 Database Normalization: GATE1999-2.7, UGCNET-June2014-III-25

https://gateoverflow.in/1485

Consider the schema $R = (S, T, U, V)$ and the dependencies $S \rightarrow T, T \rightarrow U, U \rightarrow V$ and $V \rightarrow S$. Let $R = (R1 \text{ and } R2)$ be a decomposition such that $R1 \cap R2 \neq \phi$. The decomposition is

A. not in $2NF$
B. in $2NF$ but not $3NF$
C. in $3NF$ but not in $2NF$
D. in both $2NF$ and $3NF$

gate1999 databases database-normalization normal ugcnetjune2014iii

3.5.8 Database Normalization: GATE2001-1.23, UGCNET-June2012-III-18

https://gateoverflow.in/716

Consider a schema $R(A, B, C, D)$ and functional dependencies $A \rightarrow B$ and $C \rightarrow D$. Then the decomposition of R into $R_1(A, B)$ and $R_2(C, D)$ is

A. dependency preserving and lossless join
B. lossless join but not dependency preserving
C. dependency preserving but not lossless join
D. not dependency preserving and not lossless join

gate2001 databases database-normalization normal ugcnetjune2012iii

3.5.9 Database Normalization: GATE2001-2.23

https://gateoverflow.in/741

$R(A, B, C, D)$ is a relation. Which of the following does not have a lossless join, dependency preserving $BCNF$ decomposition?

A. $A \rightarrow B, B \rightarrow CD$
B. $A \rightarrow B, B \rightarrow C, C \rightarrow D$
C. $AB \rightarrow C, C \rightarrow AD$
D. $A \rightarrow BCD$

gate2001 databases database-normalization normal

3.5.10 Database Normalization: GATE2002-16

https://gateoverflow.in/869

For relation **R=(L, M, N, O, P),** the following dependencies hold:

$M \rightarrow O$, $NO \rightarrow P$, $P \rightarrow L$ and $L \rightarrow MN$

R is decomposed into $\mathbf{R_1}$ **= (L, M, N, P)** and $\mathbf{R_2}$ **= (M, O)**.

A. Is the above decomposition a lossless-join decomposition? Explain.
B. Is the above decomposition dependency-preserving? If not, list all the dependencies that are not preserved.
C. What is the highest normal form satisfied by the above decomposition?

gate2002 databases database-normalization normal descriptive

3.5.11 Database Normalization: GATE2002-2.24

https://gateoverflow.in/854

Relation R is decomposed using a set of functional dependencies, F, and relation S is decomposed using another set of functional dependencies, G. One decomposition is definitely BCNF, the other is definitely $3NF$, but it is not known which is which. To make a guaranteed identification, which one of the following tests should be used on the decompositions? (Assume that the closures of F and G are available).

A. Dependency-preservation
B. Lossless-join
C. BCNF definition
D. $3NF$ definition

gate2002 databases database-normalization easy

3.5.12 Database Normalization: GATE2003-85

https://gateoverflow.in/968

Consider the following functional dependencies in a database.

Date_of_Birth $\rightarrow$ Age	Age $\rightarrow$ Eligibility
Name $\rightarrow$ Roll_number	Roll_number $\rightarrow$ Name
Course_number $\rightarrow$ Course_name	Course_number $\rightarrow$ Instructor
(Roll_number, Course_number) $\rightarrow$ Grade	

The relation (Roll_number, Name, Date_of_birth, Age) is

A. in second normal form but not in third normal form
B. in third normal form but not in BCNF
C. in BCNF
D. in none of the above

gate2003 databases database-normalization normal

3.5.13 Database Normalization: GATE2004-50

https://gateoverflow.in/1046

The relation scheme Student Performance (name, courseNo, rollNo, grade) has the following functional dependencies:

- name, courseNo, $\rightarrow$ grade
- rollNo, courseNo $\rightarrow$ grade
- name $\rightarrow$ rollNo
- rollNo $\rightarrow$ name

The highest normal form of this relation scheme is

A. $2NF$ B. $3NF$ C. BCNF D. $4NF$

gate2004 databases database-normalization normal

3.5.14 Database Normalization: GATE2004-IT-75

https://gateoverflow.in/3719

A relation Empdtl is defined with attributes empcode (unique), name, street, city, state and pincode. For any pincode, there is only one city and state. Also, for any given street, city and state, there is just one pincode. In normalization terms, Empdtl is a relation in

A. $1NF$ only
B. $2NF$ and hence also in $1NF$
C. $3NF$ and hence also in $2NF$ and $1NF$
D. BCNF and hence also in $3NF$, $2NF$ and $1NF$

gate2004-it databases database-normalization normal

3.5.15 Database Normalization: GATE2005-29, UGCNET-June2015-III-9

https://gateoverflow.in/1365

Which one of the following statements about normal forms is FALSE?

A. BCNF is stricter than $3NF$
B. Lossless, dependency-preserving decomposition into $3NF$ is always possible
C. Lossless, dependency-preserving decomposition into BCNF is always possible
D. Any relation with two attributes is in BCNF

gate2005 databases database-normalization easy ugcnetjune2015iii

3.5.16 Database Normalization: GATE2005-78

https://gateoverflow.in/1401

Consider a relation scheme $R = (A, B, C, D, E, H)$ on which the following functional dependencies hold: $\{A \rightarrow B$, $BC \rightarrow D$, $E \rightarrow C$, $D \rightarrow A\}$. What are the candidate keys R?

A. AE, BE B. AE, BE, DE C. AEH, BEH, BCH D. AEH, BEH, DEH

gate2005 databases database-normalization easy

3.5.17 Database Normalization: GATE2005-IT-22 https://gateoverflow.in/3767

A table has fields $Fl, F2, F3, F4, F5$ with the following functional dependencies
$F1 \rightarrow F3 \quad F2 \rightarrow F4 \quad (F1.F2) \rightarrow F5$
In terms of Normalization, this table is in

A. $1NF$ B. $2NF$ C. $3NF$ D. None of these

gate2005-it databases database-normalization easy

3.5.18 Database Normalization: GATE2007-62, UGCNET-June2014-II-47 https://gateoverflow.in/1260

Which one of the following statements is FALSE?

A. Any relation with two attributes is in BCNF
B. A relation in which every key has only one attribute is in $2NF$
C. A prime attribute can be transitively dependent on a key in a $3NF$ relation
D. A prime attribute can be transitively dependent on a key in a BCNF relation

gate2007 databases database-normalization normal ugcnetjune2014ii

3.5.19 Database Normalization: GATE2008-69 https://gateoverflow.in/492

Consider the following relational schemes for a library database:

```
Book (Title, Author, Catalog_no, Publisher, Year, Price)
Collection(Title, Author, Catalog_no)
```

with the following functional dependencies:

I. Title Author $\rightarrow$ Catalog_no
II. Catalog_no $\rightarrow$ Title Author Publisher Year
III. Publisher Title Year $\rightarrow$ Price

Assume { Author, Title } is the key for both schemes. Which of the following statements is true?

A. Both Book and Collection are in BCNF
B. Both Book and Collection are in $3NF$ only
C. Book is in $2NF$ and Collection in $3NF$
D. Both Book and Collection are in $2NF$ only

gate2008 databases database-normalization normal

3.5.20 Database Normalization: GATE2008-IT-61 https://gateoverflow.in/3371

Let $R(A, B, C, D)$ be a relational schema with the following functional dependencies :
$A \rightarrow B$, $B \rightarrow C$, $C \rightarrow D$ and $D \rightarrow B$. The decomposition of R into $(A, B), (B, C), (B, D)$

A. gives a lossless join, and is dependency preserving
B. gives a lossless join, but is not dependency preserving
C. does not give a lossless join, but is dependency preserving
D. does not give a lossless join and is not dependency preserving

gate2008-it databases database-normalization normal

3.5.21 Database Normalization: GATE2008-IT-62 https://gateoverflow.in/3372

Let $R(A, B, C, D, E, P, G)$ be a relational schema in which the following functional dependencies are known to hold: $AB \rightarrow CD, DE \rightarrow P, C \rightarrow E, P \rightarrow C$ and $B \rightarrow G$. The relational schema R is

A. in BCNF
B. in $3NF$, but not in BCNF
C. in $2NF$, but not in $3NF$
D. not in $2NF$

gate2008-it databases database-normalization normal

3.5.22 Database Normalization: GATE2009-56 https://gateoverflow.in/43474

Consider the following relational schema:

Suppliers(sid:integer, sname:string, city:string, street:string)

Parts(pid:integer, pname:string, color:string)

Catalog(sid:integer, pid:integer, cost:real)

Assume that, in the suppliers relation above, each supplier and each street within a city has unique name, and (sname, city) forms a candidate key. No other functional dependencies are implied other than those implied by primary and candidate keys. Which one of the following is TRUE about the above schema?

A. The schema is in BCNF
B. The schema is in $3NF$ but not in BCNF
C. The schema is in $2NF$ but not in $3NF$
D. The schema is not in $2NF$

gate2009 databases sql database-normalization normal

3.5.23 Database Normalization: GATE2012-2

https://gateoverflow.in/34

Which of the following is **TRUE**?

A. Every relation in 3NF is also in BCNF
B. A relation R is in 3NF if every non-prime attribute of R is fully functionally dependent on every key of R
C. Every relation in BCNF is also in 3NF
D. No relation can be in both BCNF and 3NF

gate2012 databases easy database-normalization

3.5.24 Database Normalization: GATE2014-1-30

https://gateoverflow.in/1797

Given the following two statements:

S1: Every table with two single-valued attributes is in 1NF, 2NF, 3NF and BCNF.

S2: $AB \rightarrow C$, $D \rightarrow E$, $E \rightarrow C$ is a minimal cover for the set of functional dependencies $AB \rightarrow C$, $D \rightarrow E$, $AB \rightarrow E$, $E \rightarrow C$.

Which one of the following is **CORRECT**?

A. S1 is TRUE and S2 is FALSE.
B. Both S1 and S2 are TRUE.
C. S1 is FALSE and S2 is TRUE.
D. Both S1 and S2 are FALSE.

gate2014-1 databases database-normalization normal

3.5.25 Database Normalization: GATE2016-1-21

https://gateoverflow.in/39637

Which of the following is NOT a superkey in a relational schema with attributes

V, W, X, Y, Z and primary key

$V\ Y$?

A. $VXYZ$ B. $VWXZ$ C. $VWXY$ D. $VWXYZ$

gate2016-1 databases database-normalization easy

3.5.26 Database Normalization: GATE2016-1-23

https://gateoverflow.in/39646

A database of research articles in a journal uses the following schema.

(VOLUME, NUMBER, STARTPAGE, ENDPAGE, TITLE, YEAR, PRICE)

The primary key is '(VOLUME, NUMBER, STARTPAGE, ENDPAGE)

and the following functional dependencies exist in the schema.

(VOLUME , NUMBER, STARTPAGE, ENDPAGE) $\rightarrow$ TITLE
(VOLUME, NUMBER) $\rightarrow$ YEAR
(VOLUME, NUMBER, STARTPAGE, ENDPAGE) $\rightarrow$ PRICE

The database is redesigned to use the following schemas

(VOLUME, NUMBER, STARTPAGE, ENDPAGE, TITLE, PRICE)(VOLUME, NUMBER, YEAR)

Which is the weakest normal form that the new database satisfies, but the old one does not?

A. $1NF$ B. $2NF$ C. $3NF$ D. BCNF

gate2016-1 databases database-normalization normal

3.5.27 Database Normalization: GATE2018-42

https://gateoverflow.in/204116

Consider the following four relational schemas. For each schema , all non-trivial functional dependencies alre listed, The **bolded** attributes are the respective primary keys.

Schema I: Registration(**rollno**, courses)

Field 'courses' is a set-valued attribute containing the set of courses a student has registered for.

Non-trivial functional dependency

rollno $\rightarrow$ courses

Schema II: Registration (**rollno**, **coursid**, email)

Non-trivial functional dependencies:

rollno, courseid $\rightarrow$ email

email $\rightarrow$ rollno

Schema III: Registration (**rollno**, **courseid**, marks, grade)

Non-trivial functional dependencies:

rollno, courseid, $\rightarrow$ marks, grade

marks $\rightarrow$ grade

Schema IV: Registration (**rollno**, **courseid**, credit)

Non-trivial functional dependencies:

rollno, courseid $\rightarrow$ credit

courseid $\rightarrow$ credit

Which one of the relational schemas above is in 3NF but not in BCNF?

A. Schema I B. Schema II C. Schema III D. Schema IV

gate2018 databases database-normalization normal

3.6 Deadlock (1)

3.6.1 Deadlock: GATE2004-IT-63

https://gateoverflow.in/3706

In a certain operating system, deadlock prevention is attemped using the following scheme. Each process is assigned a unique timestamp, and is restarted with the same timestamp if killed. Let P_h be the process holding a resource R, P_r be a process requesting for the same resource R, and $T(P_h)$ and $T(P_r)$ be their timestamps respectively. The decision to wait or preempt one of the processes is based on the following algorithm.

```
if T(Pr) < T(Ph) then
    kill Pr
else wait
```

Which one of the following is TRUE?

A. The scheme is deadlock-free, but not starvation-free
B. The scheme is not deadlock-free, but starvation-free
C. The scheme is neither deadlock-free nor starvation-free
D. The scheme is both deadlock-free and starvation-free

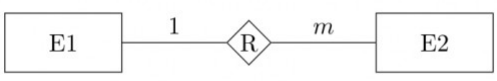
gate2004-it databases deadlock normal

3.7 Er Diagram (9)

3.7.1 Er Diagram: GATE2004-IT-73

https://gateoverflow.in/3717

Consider the following entity relationship diagram (ERD), where two entities $E1$ and $E2$ have a relation R of cardinality 1:m.



The attributes of $E1$ are $A11$, $A12$ and $A13$ where $A11$ is the key attribute. The attributes of $E2$ are $A21$, $A22$ and $A23$

where $A21$ is the key attribute and $A23$ is a multi-valued attribute. Relation R does not have any attribute. A relational database containing minimum number of tables with each table satisfying the requirements of the third normal form ($3NF$) is designed from the above ERD. The number of tables in the database is

A. 2 B. 3 C. 5 D. 4

gate2004-it databases er-diagram normal

3.7.2 Er Diagram: GATE2005-75

https://gateoverflow.in/1398

Let E_1 and E_2 be two entities in an E/R diagram with simple-valued attributes. R_1 and R_2 are two relationships between E_1 and E_2, where R_1 is one-to-many and R_2 is many-to-many. R_1 and R_2 do not have any attributes of their own. What is the minimum number of tables required to represent this situation in the relational model?

A. 2 B. 3 C. 4 D. 5

gate2005 databases er-diagram normal

3.7.3 Er Diagram: GATE2005-IT-21

https://gateoverflow.in/3766

Consider the entities 'hotel room', and 'person' with a many to many relationship 'lodging' as shown below:

If we wish to store information about the rent payment to be made by person (s) occupying different hotel rooms, then this information should appear as an attribute of

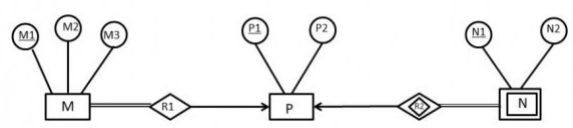
A. Person B. Hotel Room C. Lodging D. None of these

gate2005-it databases er-diagram easy

3.7.4 Er Diagram: GATE2008-82

https://gateoverflow.in/390

Consider the following ER diagram



The minimum number of tables needed to represent M, N, P, $R1$, $R2$ is

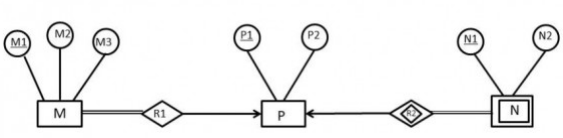
A. 2 B. 3 C. 4 D. 5

gate2008 databases er-diagram normal

3.7.5 Er Diagram: GATE2008-83

https://gateoverflow.in/87025

Consider the following ER diagram



The minimum number of tables needed to represent M, N, P, $R1$, $R2$ is

Which of the following is a correct attribute set for one of the tables for the minimum number of tables needed to represent M, N, P, $R1$, $R2$?

A. $M1, M2, M3, P1$ B. $M1, P1, N1, N2$ C. $M1, P1, N1$ D. $M1, P1$

gate2008 databases er-diagram normal

3.7.6 Er Diagram: GATE2012-14

https://gateoverflow.in/46

Given the basic ER and relational models, which of the following is **INCORRECT**?

A. An attribute of an entity can have more than one value
B. An attribute of an entity can be composite
C. In a row of a relational table, an attribute can have more than one value
D. In a row of a relational table, an attribute can have exactly one value or a NULL value

gate2012 databases normal er-diagram

3.7.7 Er Diagram: GATE2015-1-41

https://gateoverflow.in/8309

Consider an Entity-Relationship (ER) model in which entity sets E_1 and E_2 are connected by an m:n relationship R_{12}. E_1 and E_3 are connected by a 1 : n (1 on the side of E_1 and n on the side of E_3) relationship R_{13}.

E_1 has two-singled attributes a_{11} and a_{12} of which a_{11} is the key attribute. E_2 has two singled-valued attributes a_{21} and a_{22} of which a_{21} is the key attribute. E_3 has two single-valued attributes a_{31} and a_{32} of which a_{31} is the key attribute. The relationships do not have any attributes.

If a relational model is derived from the above ER model, then the minimum number of relations that would be generated if all relation are in $3NF$ is_____.

gate2015-1 databases er-diagram normal numerical-answers

3.7.8 Er Diagram: GATE2017-2-17

https://gateoverflow.in/118157

An ER model of a database consists of entity types A and B. These are connected by a relationship R which does not have its own attribute. Under which one of the following conditions, can the relational table for R be merged with that of A?

A. Relationship R is one-to-many and the participation of A in R is total
B. Relationship R is one-to-many and the participation of A in R is partial
C. Relationship R is many-to-one and the participation of A in R is total
D. Relationship R is many-to-one and the participation of A in $R$is partial

gate2017-2 databases er-diagram normal

3.7.9 Er Diagram: GATE2018-11

https://gateoverflow.in/204085

In an Entity-Relationship (ER) model, suppose R is a many-to-one relationship from entity set E1 to entity set E2. Assume that E1 and E2 participate totally in R and that the cardinality of E1 is greater than the cardinality of E2.

Which one of the following is true about R?

A. Every entity in E1 is associated with exactly one entity in E2
B. Some entity in E1 is associated with more than one entity in E2
C. Every entity in E2 is associated with exactly one entity in E1
D. Every entity in E2 is associated with at most one entity in E1

gate2018 databases er-diagram normal

3.8 Functional Dependencies (19)

3.8.1 Functional Dependencies: GATE1987-2n

https://gateoverflow.in/80609

State whether the following statements are TRUE or FALSE:

A relation r with schema (X, Y) satisfies the function dependency $X \rightarrow Y$, The tuples $\langle 1, 2\rangle$ and $\langle 2, 2\rangle$ can both be in r simultaneously.

gate1987 databases functional-dependencies

3.8.2 Functional Dependencies: GATE1988-12i

https://gateoverflow.in/94398

What are the three axioms of functional dependency for the relational databases given by Armstrong.

gate1988 normal descriptive databases functional-dependencies

3.8.3 Functional Dependencies: GATE1988-12iia

https://gateoverflow.in/94399

Using Armstrong's axioms of functional dependency derive the following rules:

$\{x \rightarrow y,\ x \rightarrow z\} \mid= x \rightarrow yz$

(Note: $x \rightarrow y$ denotes y is functionally dependent on x, $z \subseteq y$ denotes z is subset of y, and $\models$ means derives).

gate1988 easy descriptive databases functional-dependencies

3.8.4 Functional Dependencies: GATE1988-12iib

https://gateoverflow.in/94618

Using Armstrong's axioms of functional dependency derive the following rules:

$\{x \rightarrow y,\ wy \rightarrow z\} \models xw \rightarrow z$

(Note: $x \rightarrow y$ denotes y is functionally dependent on x, $z \subseteq y$ denotes z is subset of y, and $\models$ means derives).

gate1988 normal descriptive databases functional-dependencies

3.8.5 Functional Dependencies: GATE1988-12iic

https://gateoverflow.in/94619

Using Armstrong's axioms of functional dependency derive the following rules:

$\{x \rightarrow y,\ z \subset y\} \models x \rightarrow z$

(Note: $x \rightarrow y$ denotes y is functionally dependent on x, $z \subseteq y$ denotes z is subset of y, and $\models$ means derives).

gate1988 normal descriptive databases functional-dependencies

3.8.6 Functional Dependencies: GATE1990-2-iv

https://gateoverflow.in/83977

Match the pairs in the following questions:

(a)	Secondary index	(p)	Function dependency
(b)	Non-procedural query language	(q)	B-tree
(c)	Closure of a set of attributes	(r)	Domain calculus
(d)	Natural join	(s)	Relational algebraic operations

gate1990 match-the-following functional-dependencies databases

3.8.7 Functional Dependencies: GATE1990-3-ii

https://gateoverflow.in/84054

Choose the correct alternatives (More than one may be correct).

Indicate which of the following statements are true:

A relational database which is in 3NF may still have undesirable data redundancy because there may exist:

A. Transitive functional dependencies
B. Non-trivial functional dependencies involving prime attributes on the right-side.
C. Non-trivial functional dependencies involving prime attributes only on the left-side.
D. Non-trivial functional dependencies involving only prime attributes.

gate1990 normal databases functional-dependencies

3.8.8 Functional Dependencies: GATE2000-2.24

https://gateoverflow.in/671

Given the following relation instance.

X	Y	Z
1	4	2
1	5	3
1	6	3
3	2	2

Which of the following functional dependencies are satisfied by the instance?

A. $XY \rightarrow Z$ and $Z \rightarrow Y$

B. $YZ \rightarrow X$ and $Y \rightarrow Z$

C. $YZ \rightarrow X$ and $X \rightarrow Z$ D. $XZ \rightarrow Y$ and $Y \rightarrow X$

gate2000 databases functional-dependencies easy

3.8.9 Functional Dependencies: GATE2002-1.19

https://gateoverflow.in/824

Relation R with an associated set of functional dependencies, F, is decomposed into BCNF. The redundancy (arising out of functional dependencies) in the resulting set of relations is

A. Zero
B. More than zero but less than that of an equivalent $3NF$ decomposition
C. Proportional to the size of F^+
D. Indeterminate

gate2002 databases functional-dependencies database-normalization normal

3.8.10 Functional Dependencies: GATE2002-2.25

https://gateoverflow.in/855

From the following instance of a relation schema $R(A, B, C)$, we can conclude that:

A	B	C
1	1	1
1	1	0
2	3	2
2	3	2

A. A functionally determines B and B functionally determines C
B. A functionally determines B and B does not functionally determine C
C. B does not functionally determine C
D. A does not functionally determine B and B does not functionally determine C

gate2002 databases functional-dependencies

3.8.11 Functional Dependencies: GATE2005-IT-70

https://gateoverflow.in/3833

In a schema with attributes A, B, C, D and E following set of functional dependencies are given

- $A \rightarrow B$
- $A \rightarrow C$
- $CD \rightarrow E$
- $B \rightarrow D$
- $E \rightarrow A$

Which of the following functional dependencies is NOT implied by the above set?

A. $CD \rightarrow AC$ B. $BD \rightarrow CD$ C. $BC \rightarrow CD$ D. $AC \rightarrow BC$

gate2005-it databases functional-dependencies normal

3.8.12 Functional Dependencies: GATE2006-70

https://gateoverflow.in/1848

The following functional dependencies are given:

$AB \rightarrow CD, AF \rightarrow D, DE \rightarrow F, C \rightarrow G, F \rightarrow E, G \rightarrow A$

Which one of the following options is false?

A. $\{CF\}^* = \{ACDEFG\}$ B. $\{BG\}^* = \{ABCDG\}$
C. $\{AF\}^* = \{ACDEFG\}$ D. $\{AB\}^* = \{ABCDG\}$

gate2006 databases functional-dependencies normal

3.8.13 Functional Dependencies: GATE2006-IT-60

https://gateoverflow.in/3604

Consider a relation R with five attributes V, W, X, Y, and Z. The following functional dependencies hold:

VY→ W, WX → Z, and ZY → V.

Which of the following is a candidate key for R?

A. VXZ B. VXY C. VWXY D. VWXYZ

gate2006-it databases functional-dependencies normal

3.8.14 Functional Dependencies: GATE2013-54

https://gateoverflow.in/1558

Relation R has eight attributes ABCDEFGH. Fields of R contain only atomic values. F= $\{\mathrm{CH}\rightarrow\mathrm{G},\ \mathrm{A}\rightarrow\mathrm{BC},\ \mathrm{B}\rightarrow\mathrm{CFH},\ \mathrm{E}\rightarrow\mathrm{A},\ \mathrm{F}\rightarrow\mathrm{EG}\}$ is a set of functional dependencies (FDs) so that F^+ is exactly the set of FDs that hold for R.

How many candidate keys does the relation R have?

A. 3 B. 4 C. 5 D. 6

gate2013 databases functional-dependencies database-normalization normal

3.8.15 Functional Dependencies: GATE2013-55

https://gateoverflow.in/43290

Relation R has eight attributes ABCDEFGH. Fields of R contain only atomic values. $F = \{\mathrm{CH} \rightarrow \mathrm{G},\ \mathrm{A} \rightarrow \mathrm{BC},\ \mathrm{B} \rightarrow \mathrm{CFH},\ \mathrm{E} \rightarrow \mathrm{A},\ \mathrm{F} \rightarrow \mathrm{EG}\}$ is a set of functional dependencies (FDs) so that F^+ is exactly the set of FDs that hold for R.

The relation R is

A. in $1NF$, but not in $2NF$.
B. in $2NF$, but not in $3NF$.
C. in $3NF$, but not in BCNF.
D. in BCNF.

gate2013 databases functional-dependencies database-normalization normal

3.8.16 Functional Dependencies: GATE2014-1-21

https://gateoverflow.in/1788

Consider the relation scheme $R = (E, F, G, H, I, J, K, L, M, N)$ and the set of functional dependencies

$$\{\{E, F\} \rightarrow \{G\}, \{F\} \rightarrow \{I, J\}, \{E, H\} \rightarrow \{K, L\}, \{K\} \rightarrow \{M\}, \{L\} \rightarrow \{N\}\}$$

on R. What is the key for R?

A. $\{E, F\}$ B. $\{E, F, H\}$ C. $\{E, F, H, K, L\}$ D. $\{E\}$

gate2014-1 databases functional-dependencies normal

3.8.17 Functional Dependencies: GATE2015-3-20

https://gateoverflow.in/8420

Consider the relation $X(P, Q, R, S, T, U)$ with the following set of functional dependencies

$F = \{\ \{P, R\} \rightarrow \{S, T\},\ \{P, S, U\} \rightarrow \{Q, R\}\ \}$

Which of the following is the trivial functional dependency in F^+, where F^+ is closure to F?

A. $\{P, R\} \rightarrow \{S, T\}$
B. $\{P, R\} \rightarrow \{R, T\}$
C. $\{P, S\} \rightarrow \{S\}$
D. $\{P, S, U\} \rightarrow \{Q\}$

gate2015-3 databases functional-dependencies easy

3.8.18 Functional Dependencies: GATE2017-1-16

https://gateoverflow.in/118296

The following functional dependencies hold true for the relational schema $R\{V, W, X, Y, Z\}$:

$\mathrm{V} \rightarrow \mathrm{W}$
$\mathrm{VW} \rightarrow \mathrm{X}$
$\mathrm{Y} \rightarrow \mathrm{VX}$
$\mathrm{Y} \rightarrow \mathrm{Z}$

Which of the following is irreducible equivalent for this set of functional dependencies?

A. $\mathrm{V} \rightarrow \mathrm{W}$
$\mathrm{V} \rightarrow \mathrm{X}$
$\mathrm{Y} \rightarrow \mathrm{V}$
$\mathrm{Y} \rightarrow \mathrm{Z}$

B. $\mathrm{V} \rightarrow \mathrm{W}$
$\mathrm{W} \rightarrow \mathrm{X}$

$Y \rightarrow V$
$Y \rightarrow Z$

C. $V \rightarrow W$
$V \rightarrow X$
$Y \rightarrow V$
$Y \rightarrow X$
$Y \rightarrow Z$

D. $V \rightarrow W$
$W \rightarrow X$
$Y \rightarrow V$
$Y \rightarrow X$
$Y \rightarrow Z$

gate2017-1 databases functional-dependencies normal

3.8.19 Functional Dependencies: GATE2019-32

https://gateoverflow.in/302816

Let the set of functional dependencies $F = \{QR \rightarrow S,\ R \rightarrow P,\ S \rightarrow Q\}$ hold on a relation schema $X = (PQRS)$. X is not in BCNF. Suppose X is decomposed into two schemas Y and Z, where $Y = (PR)$ and $Z = (QRS)$.

Consider the two statements given below.

I. Both Y and Z are in BCNF
II. Decomposition of X into Y and Z is dependency preserving and lossless

Which of the above statements is/are correct?

A. Both I and II B. I only C. II only D. Neither I nor II

gate2019 databases functional-dependencies

3.9 Indexing (10)

3.9.1 Indexing: GATE1989-4-xiv

https://gateoverflow.in/88228

Provide short answers to the following questions:

For secondary key processing which of the following file organizations is preferred? Give a one line justification:

A. Indexed sequential file organization.
B. Two-way linked list.
C. Inverted file organization.
D. Sequential file organization.

gate1989 normal databases indexing

3.9.2 Indexing: GATE1990-10b

https://gateoverflow.in/85691

One giga bytes of data are to be organized as an indexed-sequential file with a uniform blocking factor 8. Assuming a block size of 1 Kilo bytes and a block refrencing pointer size of 32 bits, find out the number of levels of indexing that would be required and the size of the index at each level. Determine also the size of the master index. The referencing capability (fanout ratio) per block of index storage may be considered to be 32.

gate1990 descriptive databases indexing

3.9.3 Indexing: GATE1993-14

https://gateoverflow.in/2311

An ISAM (indexed sequential) file consists of records of size $64\ bytes$ each, including key field of size $14\ bytes$. An address of a disk block takes $2\ bytes$. If the disk block size is $512\ bytes$ and there are $16\ K$ records, compute the size of the data and index areas in terms of number blocks. How many levels of tree do you have for the index?

gate1993 databases indexing normal

3.9.4 Indexing: GATE1998-1.35

https://gateoverflow.in/1672

There are five records in a database.

Name	Age	Occupation	Category
Rama	27	CON	A
Abdul	22	ENG	A
Jennifer	28	DOC	B
Maya	32	SER	D
Dev	24	MUS	C

There is an index file associated with this and it contains the values $1, 3, 2, 5$ and 4. Which one of the fields is the index built from?

A. Age B. Name C. Occupation D. Category

gate1998 databases indexing normal

3.9.5 Indexing: GATE2002-2.22

https://gateoverflow.in/852

In the index allocation scheme of blocks to a file, the maximum possible size of the file depends on

A. the size of the blocks, and the size of the address of the blocks.
B. the number of blocks used for the index, and the size of the blocks.
C. the size of the blocks, the number of blocks used for the index, and the size of the address of the blocks.
D. None of the above

gate2002 databases indexing normal

3.9.6 Indexing: GATE2008-16, ISRO2016-60

https://gateoverflow.in/414

A clustering index is defined on the fields which are of type

A. non-key and ordering
B. non-key and non-ordering
C. key and ordering
D. key and non-ordering

gate2008 easy databases indexing isro2016

3.9.7 Indexing: GATE2008-70

https://gateoverflow.in/259

Consider a file of 16384 records. Each record is 32 $bytes$ long and its key field is of size 6 $bytes$. The file is ordered on a non-key field, and the file organization is unspanned. The file is stored in a file system with block size 1024 $bytes$, and the size of a block pointer is 10 $bytes$. If the secondary index is built on the key field of the file, and a multi-level index scheme is used to store the secondary index, the number of first-level and second-level blocks in the multi-level index are respectively

A. 8 and 0 B. 128 and 6 C. 256 and 4 D. 512 and 5

gate2008 databases indexing normal

3.9.8 Indexing: GATE2011-39

https://gateoverflow.in/2141

Consider a relational table r with sufficient number of records, having attributes $A_1, A_2, \ldots, A_n$ and let $1 \leq p \leq n$. Two queries $Q1$ and $Q2$ are given below.

- $Q1 : \pi_{A_1,\ldots,A_p}\left(\sigma_{A_p=c}(r)\right)$ where c is a constant
- $Q2 : \pi_{A_1,\ldots,A_p}\left(\sigma_{c_1 \leq A_p \leq c_2}(r)\right)$ where c_1 and c_2 are constants.

The database can be configured to do ordered indexing on A_p or hashing on A_p. Which of the following statements is **TRUE**?

A. Ordered indexing will always outperform hashing for both queries
B. Hashing will always outperform ordered indexing for both queries
C. Hashing will outperform ordered indexing on $Q1$, but not on $Q2$
D. Hashing will outperform ordered indexing on $Q2$, but not on $Q1$

gate2011 databases indexing normal

3.9.9 Indexing: GATE2013-15

https://gateoverflow.in/1437

An index is clustered, if

A. it is on a set of fields that form a candidate key
B. it is on a set of fields that include the primary key
C. the data records of the file are organized in the same order as the data entries of the index
D. the data records of the file are organized not in the same order as the data entries of the index

gate2013 databases indexing normal

3.9.10 Indexing: GATE2015-1-24

https://gateoverflow.in/8222

A file is organized so that the ordering of the data records is the same as or close to the ordering of data entries in some index. Than that index is called

A. Dense B. Sparse C. Clustered D. Unclustered

gate2015-1 databases indexing easy

3.10 Joins (7)

3.10.1 Joins: GATE2004-14

https://gateoverflow.in/1011

Consider the following relation schema pertaining to a students database:

- Students (rollno, name, address)
- Enroll (rollno, courseno, coursename)

where the primary keys are shown underlined. The number of tuples in the student and Enroll tables are 120 and 8 respectively. What are the maximum and minimum number of tuples that can be present in (Student * Enroll), where '*' denotes natural join?

A. $8, 8$ B. $120, 8$ C. $960, 8$ D. $960, 120$

gate2004 databases easy joins natural-join

3.10.2 Joins: GATE2005-IT-82a

https://gateoverflow.in/3847

A database table T_1 has 2000 records and occupies 80 disk blocks. Another table T_2 has 400 records and occupies 20 disk blocks. These two tables have to be joined as per a specified join condition that needs to be evaluated for every pair of records from these two tables. The memory buffer space available can hold exactly one block of records for T_1 and one block of records for T_2 simultaneously at any point in time. No index is available on either table.

If Nested-loop join algorithm is employed to perform the join, with the most appropriate choice of table to be used in outer loop, the number of block accesses required for reading the data are

A. 800000 B. 40080 C. 32020 D. 100

gate2005-it databases normal joins

3.10.3 Joins: GATE2005-IT-82b

https://gateoverflow.in/3848

A database table T_1 has 2000 records and occupies 80 disk blocks. Another table T_2 has 400 records and occupies 20 disk blocks. These two tables have to be joined as per a specified join condition that needs to be evaluated for every pair of records from these two tables. The memory buffer space available can hold exactly one block of records for T_1 and one block of records for T_2 simultaneously at any point in time. No index is available on either table.

If, instead of Nested-loop join, Block nested-loop join is used, again with the most appropriate choice of table in the outer loop, the reduction in number of block accesses required for reading the data will be

A. 0 B. 30400 C. 38400 D. 798400

gate2005-it databases normal joins

3.10.4 Joins: GATE2006-IT-14

https://gateoverflow.in/3553

Consider the relations $r_1(P, Q, R)$ and $r_2(R, S, T)$ with primary keys P and R respectively. The relation r_1 contains 2000 tuples and r_2 contains 2500 tuples. The maximum size of the join $r_1 \bowtie r_2$ is :

A. 2000 B. 2500 C. 4500 D. 5000

gate2006-it databases joins natural-join normal

3.10.5 Joins: GATE2007-IT-68

https://gateoverflow.in/3513

Consider the following relation schemas :

- b-Schema = (b-name, b-city, assets)
- a-Schema = (a-num, b-name, bal)
- d-Schema = (c-name, a-number)

Let branch, account and depositor be respectively instances of the above schemas. Assume that account and depositor relations are much bigger than the branch relation.

Consider the following query:

$\Pi_{\text{c-name}} (\sigma_{\text{b-city = "Agra"} \wedge \text{bal} < 0} (\text{branch} \bowtie (\text{account} \bowtie \text{depositor})$

Which one of the following queries is the most efficient version of the above query ?

A. $\Pi_{\text{c-name}} (\sigma_{\text{bal} < 0} (\sigma_{\text{b-city = "Agra"}} \text{branch} \bowtie \text{account}) \bowtie \text{depositor})$
B. $\Pi_{\text{c-name}} (\sigma_{\text{b-city = "Agra"}} \text{branch} \bowtie (\sigma_{\text{bal} < 0} \text{account} \bowtie \text{depositor}))$
C. $\Pi_{\text{c-name}} ((\sigma_{\text{b-city = "Agra"}} \text{branch} \bowtie \sigma_{\text{b-city = "Agra"} \wedge \text{bal} < 0} \text{account}) \bowtie \text{depositor})$
D. $\Pi_{\text{c-name}} (\sigma_{\text{b-city = "Agra"}} \text{branch} \bowtie (\sigma_{\text{b-city = "Agra"} \wedge \text{bal} < 0} \text{account} \bowtie \text{depositor}))$

gate2007-it databases joins relational-algebra normal

3.10.6 Joins: GATE2012-50

https://gateoverflow.in/2180

Consider the following relations A, B and C :

A

ID	Name	Age
12	Arun	60
15	Shreya	24
99	Rohit	11

B

ID	Name	Age
15	Shreya	24
25	Hari	40
98	Rohit	20
99	Rohit	11

C

ID	Phone	Area
10	2200	02
99	2100	01

How many tuples does the result of the following relational algebra expression contain? Assume that the schema of $A \cup B$ is the same as that of A.

$$(A \cup B) \bowtie_{A.Id>40 \vee C.Id<15} C$$

A. 7 B. 4 C. 5 D. 9

gate2012 databases joins normal

3.10.7 Joins: GATE2014-2-30

https://gateoverflow.in/1989

Consider a join (relation algebra) between relations $r(R)$ and $s(S)$ using the nested loop method. There are 3 buffers each of size equal to disk block size, out of which one buffer is reserved for intermediate results. Assuming size(r(R)) <size(s(S)) , the join will have fewer number of disk block accesses if

A. relation $r(R)$ is in the outer loop.
B. relation $s(S)$ is in the outer loop.
C. join selection factor between $r(R)$ and $s(S)$ is more than 0.5.
D. join selection factor between $r(R)$ and $s(S)$ is less than 0.5.

gate2014-2 databases normal joins

3.11 Multivalued Dependency 4nf (1)

3.11.1 Multivalued Dependency 4nf: GATE2007-IT-67

https://gateoverflow.in/3512

Consider the following implications relating to functional and multivalued dependencies given below, which may or may not be correct.

i. if $A \rightarrow\rightarrow B$ and $A \rightarrow\rightarrow C$ then $A \rightarrow\rightarrow BC$
ii. if $A \rightarrow B$ and $A \rightarrow C$ then $A \rightarrow\rightarrow BC$
iii. if $A \rightarrow\rightarrow BC$ and $A \rightarrow B$ then $A \rightarrow C$
iv. if $A \rightarrow BC$ and $A \rightarrow B$ then $A \rightarrow\rightarrow C$

i. If A$\rightarrow\rightarrow$ B and A $\rightarrow\rightarrow$ C then A $\rightarrow$ BC
ii. If A $\rightarrow$ B and A $\rightarrow$ C then A $\rightarrow\rightarrow$ BC
iii. If A $\rightarrow\rightarrow$ BC and A $\rightarrow$ B then A $\rightarrow$ C
iv. If A $\rightarrow$ BC and A $\rightarrow$ B then A $\rightarrow\rightarrow$ C

Exactly how many of the above implications are valid?

A. 0 B. 1 C. 2 D. 3

gate2007-it databases functional-dependencies multivalued-dependency-4nf normal

3.12 Natural Join (3)

3.12.1 Natural Join: GATE2005-30

https://gateoverflow.in/1366

Let r be a relation instance with schema R = (A, B, C, D). We define $r_1 = \pi_{A,B,C}(R)$ and $r_2 = \pi_{A,D}(r)$. Let $s = r_1 * r_2$ where $*$ denotes natural join. Given that the decomposition of r into r_1 and r_2 is lossy, which one of the following is TRUE?

A. $s \subset r$ B. $r \cup s = r$ C. $r \subset s$ D. $r * s = s$

gate2005 databases relational-algebra natural-join normal

3.12.2 Natural Join: GATE2010-43

https://gateoverflow.in/2344

The following functional dependencies hold for relations $R(A, B, C)$ and $S(B, D, E)$.

- $B \rightarrow A$
- $A \rightarrow C$

The relation R contains 200 tuples and the relation S contains 100 tuples. What is the maximum number of tuples possible in the natural join $R \bowtie S$?

A. 100 B. 200 C. 300 D. 2000

gate2010 databases normal natural-join functional-dependencies

3.12.3 Natural Join: GATE2015-2-32

https://gateoverflow.in/8151

Consider two relations $R_1(A, B)$ with the tuples $(1, 5), (3, 7)$ and $R_2(A, C) = (1, 7), (4, 9)$. Assume that $R(A, B, C)$ is the full natural outer join of R_1 and R_2. Consider the following tuples of the form (A,B,C):
$a = (1, 5, null), b = (1, null, 7), c = (3, null, 9), d = (4, 7, null), e = (1, 5, 7), f = (3, 7, null), g = (4, null, 9)$.
Which one of the following statements is correct?

A. R contains a, b, e, f, g but not c, d.
B. R contains all a, b, c, d, e, f, g.
C. R contains e, f, g but not a, b.
D. R contains e but not f, g.

gate2015-2 databases normal natural-join

3.13 Referential Integrity (3)

3.13.1 Referential Integrity: GATE1997-6.10, ISRO2016-54

https://gateoverflow.in/2266

Let $R(a, b, c)$ and $S(d, e, f)$ be two relations in which d is the foreign key of S that refers to the primary key of R.

Consider the following four operations R and S

I. Insert into R
II. Insert into S
III. Delete from R
IV. Delete from S

Which of the following can cause violation of the referential integrity constraint above?

A. Both I and IV B. Both II and III C. All of these D. None of these

gate1997 databases referential-integrity easy isro2016

3.13.2 Referential Integrity: GATE2005-76

https://gateoverflow.in/1399

The following table has two attributes A and C where A is the primary key and C is the foreign key referencing A with on-delete cascade.

A	C
2	4
3	4
4	3
5	2
7	2
9	5
6	4

The set of all tuples that must be additionally deleted to preserve referential integrity when the tuple $(2,4)$ is deleted is:

A. $(3,4)$ and $(6,4)$
B. $(5,2)$ and $(7,2)$
C. $(5,2)$, $(7,2)$ and $(9,5)$
D. $(3,4)$, $(4,3)$ and $(6,4)$

gate2005 databases referential-integrity normal

3.13.3 Referential Integrity: GATE2017-2-19

https://gateoverflow.in/118236

Consider the following tables $T1$ and $T2$.

$T1$

P	Q
2	2
3	8
7	3
5	8
6	9
8	5
9	8

$T2$

R	S
2	2
8	3
3	2
9	7
5	7
7	2

In table $T1$ **P** is the primary key and **Q** is the foreign key referencing **R** in table $T2$ with on-delete cascade and on-update cascade. In table $T2$, **R** is the primary key and **S** is the foreign key referencing **P** in table $T1$ with on-delete set NULL and on-update cascade. In order to delete record $\langle 3,8 \rangle$ from the table $T1$, the number of additional records that need to be deleted from table $T1$ is _____

gate2017-2 databases numerical-answers referential-integrity normal

3.14 Relational Algebra (26)

3.14.1 Relational Algebra: GATE1992-13b

https://gateoverflow.in/43581

Suppose we have a database consisting of the following three relations:

FREQUENTS	(CUSTOMER, HOTEL)
SERVES	(HOTEL, SNACKS)
LIKES	(CUSTOMER, SNACKS)

The first indicates the hotels each customer visits, the second tells which snacks each hotel serves and last indicates which snacks are liked by each customer. Express the following query in relational algebra:

Print the hotels the serve the snack that customer Rama likes.

gate1992 databases relational-algebra normal

3.14.2 Relational Algebra: GATE1994-13

https://gateoverflow.in/2509

Consider the following relational schema:

- COURSES (cno, cname)
- STUDENTS (rollno, sname, age, year)
- REGISTERED_FOR (cno, rollno)

The underlined attributes indicate the primary keys for the relations. The 'year' attribute for the STUDENTS relation indicates the year in which the student is currently studying (First year, Second year etc.)

a. Write a relational algebra query to print the roll number of students who have registered for cno 322.
b. Write a SQL query to print the age and year of the youngest student in each year.

gate1994 databases relational-algebra sql normal

3.14.3 Relational Algebra: GATE1995-27

https://gateoverflow.in/2666

Consider the relation scheme.

AUTHOR	(ANAME, INSTITUTION, ACITY, AGE)
PUBLISHER	(PNAME, PCITY)
BOOK	(TITLE, ANAME, PNAME)

Express the following queries using (one or more of) SELECT, PROJECT, JOIN and DIVIDE operations.

a. Get the names of all publishers.
b. Get values of all attributes of all authors who have published a book for the publisher with PNAME='TECHNICAL PUBLISHERS'.
c. Get the names of all authors who have published a book for any publisher located in Madras

gate1995 databases relational-algebra normal

3.14.4 Relational Algebra: GATE1996-27

https://gateoverflow.in/2779

A library relational database system uses the following schema

- USERS (User#, User Name, Home Town)
- BOOKS (Book#, Book Title, Author Name)
- ISSUED (Book#, User#, Date)

Explain in one English sentence, what each of the following relational algebra queries is designed to determine

a. $\sigma_{\text{User\#}=6}\left(\pi_{\text{User\#, Book Title}}\left((\text{USERS} \bowtie \text{ISSUED}) \bowtie \text{BOOKS}\right)\right)$
b. $\pi_{\text{Author Name}}\left(\text{BOOKS} \bowtie \sigma_{\text{Home Town}=\text{Delhi}}\left(\text{USERS} \bowtie \text{ISSUED}\right)\right)$

gate1996 databases relational-algebra normal

3.14.5 Relational Algebra: GATE1997-76-a

https://gateoverflow.in/19838

Consider the following relational database schema:

- EMP (eno name, age)
- PROJ (pno name)
- INVOLVED (eno, pno)

EMP contains information about employees. PROJ about projects and involved about which employees involved in which projects. The underlined attributes are the primary keys for the respective relations.

What is the relational algebra expression containing one or more of $\{\sigma, \pi, \times, \rho, -\}$ which is equivalent to SQL query.

```
select eno  from EMP|INVOLVED where EMP.eno=INVOLVED.eno  and INVOLVED.pno=3
```

gate1997 databases sql relational-algebra

3.14.6 Relational Algebra: GATE1998-1.33

https://gateoverflow.in/1670

Given two union compatible relations $R_1(A, B)$ and $R_2(C, D)$, what is the result of the operation $R_1 \bowtie_{A=C \wedge B=D} R_2$?

A. $R_1 \cup R_2$ B. $R_1 \times R_2$ C. $R_1 - R_2$ D. $R_1 \cap R_2$

gate1998 normal relational-algebra

3.14.7 Relational Algebra: GATE1998-27

https://gateoverflow.in/1742

Consider the following relational database schemes:

- COURSES (Cno, Name)
- PRE_REQ(Cno, Pre_Cno)
- COMPLETED (Student_no, Cno)

COURSES gives the number and name of all the available courses.

PRE_REQ gives the information about which courses are pre-requisites for a given course.

COMPLETED indicates what courses have been completed by students

Express the following using relational algebra:

List all the courses for which a student with Student_no 2310 has completed all the pre-requisites.

gate1998 databases relational-algebra normal

3.14.8 Relational Algebra: GATE1999-1.18, ISRO2016-53

https://gateoverflow.in/1471

Consider the join of a relation R with a relation S. If R has m tuples and S has n tuples then the maximum and minimum sizes of the join respectively are

A. $m+n$ and 0 B. mn and 0 C. $m+n$ and $|m-n|$ D. mn and $m+n$

gate1999 databases relational-algebra easy isro2016

3.14.9 Relational Algebra: GATE2000-1.23, ISRO2016-57

https://gateoverflow.in/647

Given the relations

- employee (name, salary, dept-no), and
- department (dept-no, dept-name,address),

Which of the following queries cannot be expressed using the basic relational algebra operations $(\sigma, \pi, \times, \bowtie, \cup, \cap, -)$?

A. Department address of every employee
B. Employees whose name is the same as their department name
C. The sum of all employees' salaries
D. All employees of a given department

gate2000 databases relational-algebra easy isro2016

3.14.10 Relational Algebra: GATE2001-1.24

https://gateoverflow.in/717

Suppose the adjacency relation of vertices in a graph is represented in a table Adj (X,Y). Which of the following queries cannot be expressed by a relational algebra expression of constant length?

A. List all vertices adjacent to a given vertex
B. List all vertices which have self loops
C. List all vertices which belong to cycles of less than three vertices
D. List all vertices reachable from a given vertex

D. This is most hard part. Here we need to find closure of vertices. This willl need kind of loop. If the graph is like skewed tree, our query must loop for O(N) times. We can't do with constant length query here.

gate2001 databases relational-algebra normal

3.14.11 Relational Algebra: GATE2001-1.25

https://gateoverflow.in/718

Let r and s be two relations over the relation schemes R and S respectively, and let A be an attribute in R. The relational algebra expression $\sigma_{A=a}(r \bowtie s)$ is always equal to

A. $\sigma_{A=a}(r)$
B. r
C. $\sigma_{A=a}(r) \bowtie s$
D. None of the above

gate2001 databases relational-algebra

3.14.12 Relational Algebra: GATE2002-15

https://gateoverflow.in/868

A university placement center maintains a relational database of companies that interview students on campus and make job offers to those successful in the interview. The schema of the database is given below:

COMPANY(<u>cname</u>, clocation)	STUDENT(<u>srollno</u>, sname, sdegree)
INTERVIEW(<u>cname, srollno</u>, idate)	OFFER(<u>cname, srollno</u>, osalary)

The COMPANY relation gives the name and location of the company. The STUDENT relation gives the student's roll number, name and the degree program for which the student is registered in the university. The INTERVIEW relation gives the date on which a student is interviewed by a company. The OFFER relation gives the salary offered to a student who is successful in a company's interview. The key for each relation is indicated by the underlined attributes

a. Write a **relational algebra** expressions (using only the operators $\bowtie, \sigma, \pi, \cup, -$) for the following queries.
   i. List the $rollnumbers$ and $names$ of students who attended at least one interview but did not receive any job offer.
   ii. List the $rollnumbers$ and $names$ of students who went for interviews and received job offers from $every$ company with which they interviewed.
b. Write an SQL query to list, for each degree program in which more than $five$ students were offered jobs, the name of the degree and the average offered salary of students in this degree program.

gate2002 databases normal descriptive relational-algebra sql

3.14.13 Relational Algebra: GATE2003-30

https://gateoverflow.in/920

Consider the following SQL query

Select distinct $a_1, a_2, \ldots, a_n$

from $r_1, r_2, \ldots, r_m$

where P

For an arbitrary predicate P, this query is equivalent to which of the following relational algebra expressions?

A. $\Pi_{a_1,a_2,\ldots a_n}\sigma_p(r_1 \times r_2 \times \cdots \times r_m)$
B. $\Pi_{a_1,a_2,\ldots a_n}\sigma_p(r_1 \bowtie r_2 \bowtie \cdots \bowtie r_m)$
C. $\Pi_{a_1,a_2,\ldots a_n}\sigma_p(r_1 \cup r_2 \cup \cdots \cup r_m)$
D. $\Pi_{a_1,a_2,\ldots a_n}\sigma_p(r_1 \cap r_2 \cap \cdots \cap r_m)$

gate2003 databases relational-algebra normal

3.14.14 Relational Algebra: GATE2004-51

https://gateoverflow.in/1047

Consider the relation Student (name, sex, marks), where the primary key is shown underlined, pertaining to students in a class that has at least one boy and one girl. What does the following relational algebra expression produce? (Note: ρ is the rename operator).

$\pi_{name}\{\sigma_{sex=female}(\text{Student})\} - \pi_{name}(\text{Student}_{(sex=female \wedge x=male \wedge marks \leq m)}\rho_{n,x,m}(\text{Student}))$

A. names of girl students with the highest marks
B. names of girl students with more marks than some boy student
C. names of girl students with marks not less than some boy student
D. names of girl students with more marks than all the boy students

gate2004 databases relational-algebra normal

3.14.15 Relational Algebra: GATE2005-IT-68

https://gateoverflow.in/3831

A table 'student' with schema (roll, name, hostel, marks), and another table 'hobby' with schema (roll, hobbyname) contains records as shown below:

Table: student

Roll	Name	Hostel	Marks
1798	Manoj Rathor	7	95
2154	Soumic Banerjee	5	68
2369	Gumma Reddy	7	86
2581	Pradeep pendse	6	92
2643	Suhas Kulkarni	5	78
2711	Nitin Kadam	8	72
2872	Kiran Vora	5	92
2926	Manoj Kunkalikar	5	94
2959	Hemant Karkhanis	7	88
3125	Rajesh Doshi	5	82

Table: hobby

Roll	Hobby Name
1798	chess
1798	music
2154	music
2369	swimming
2581	cricket
2643	chess
2643	hockey
2711	volleyball
2872	football
2926	cricket
2959	photography
3125	music
3125	chess

The following SQL query is executed on the above tables:

```
select hostel
from student natural join hobby
where marks >= 75 and roll between 2000 and 3000;
```

Relations S and H with the same schema as those of these two tables respectively contain the same information as tuples. A new relation S' is obtained by the following relational algebra operation:

$$S' = \Pi_{\text{hostel}}((\sigma_{s.roll=H.roll}(\sigma_{marks>75 \text{ and } roll>2000 \text{ and } roll<3000}(S)) \times (H))$$

The difference between the number of rows output by the SQL statement and the number of tuples in S' is

A. 6 B. 4 C. 2 D. 0

gate2005-it databases sql relational-algebra normal

3.14.16 Relational Algebra: GATE2007-59

https://gateoverflow.in/2428

Information about a collection of students is given by the relation studInfo(studId, name, sex). The relation enroll(studId, courseId) gives which student has enrolled for (or taken) what course(s). Assume that every course is taken by at least one male and at least one female student. What does the following relational algebra expression represent?

$\pi_{courceId}((\pi_{\text{studId}}(\sigma_{sex=\text{"}female\text{"}}(\text{studInfo})) \times \pi_{courseId}(\text{enroll})) - \text{enroll})$

A. Courses in which all the female students are enrolled.
B. Courses in which a proper subset of female students are enrolled.
C. Courses in which only male students are enrolled.
D. None of the above

gate2007 databases relational-algebra normal

3.14.17 Relational Algebra: GATE2008-68

https://gateoverflow.in/491

Let R and S be two relations with the following schema

$R(\underline{P,Q}, R1, R2, R3)$

$S(\underline{P,Q}, S1, S2)$

where $\{P, Q\}$ is the key for both schemas. Which of the following queries are equivalent?

I. $\Pi_P (R \bowtie S)$
II. $\Pi_P (R) \bowtie \Pi_P (S)$
III. $\Pi_P (\Pi_{P,Q} (R) \cap \Pi_{P,Q} (S))$
IV. $\Pi_P (\Pi_{P,Q} (R) - (\Pi_{P,Q} (R) - \Pi_{P,Q} (S)))$

A. Only I and II B. Only I and III C. Only I, II and III D. Only I, III and IV

gate2008 databases relational-algebra normal

3.14.18 Relational Algebra: GATE2012-43

https://gateoverflow.in/2151

Suppose $R_1(\underline{A}, B)$ and $R_2(\underline{C}, D)$ are two relation schemas. Let r_1 and r_2 be the corresponding relation instances. B is a foreign key that refers to C in R_2. If data in r_1 and r_2 satisfy referential integrity constraints, which of the following is **ALWAYS TRUE**?

A. $\prod_B(r_1) - \prod_C(r_2) = \varnothing$
B. $\prod_C(r_2) - \prod_B(r_1) = \varnothing$
C. $\prod_B(r_1) = \prod_C(r_2)$
D. $\prod_B(r_1) - \prod_C(r_2) \neq \varnothing$

gate2012 databases relational-algebra normal

3.14.19 Relational Algebra: GATE2014-3-21

https://gateoverflow.in/2055

What is the optimized version of the relation algebra expression $\pi_{A1}(\pi_{A2}(\sigma_{F1}(\sigma_{F2}(r))))$, where $A1, A2$ are sets of attributes in r with $A1 \subset A2$ and $F1, F2$ are Boolean expressions based on the attributes in r?

A. $\pi_{A1}(\sigma_{(F1 \wedge F2)}(r))$
B. $\pi_{A1}(\sigma_{(F1 \vee F2)}(r))$
C. $\pi_{A2}(\sigma_{(F1 \wedge F2)}(r))$
D. $\pi_{A2}(\sigma_{(F1 \vee F2)}(r))$

gate2014-3 databases relational-algebra easy

3.14.20 Relational Algebra: GATE2014-3-30

https://gateoverflow.in/2064

Consider the relational schema given below, where **eId** of the relation **dependent** is a foreign key referring to **empId** of the relation **employee**. Assume that every employee has at least one associated dependent in the **dependent** relation.

employee (empId, empName, empAge)

dependent (depId, eId, depName, depAge)

Consider the following relational algebra query:

$\Pi_{empId}\ (employee) - \Pi_{empId}\ (employee \bowtie_{(empId=eID) \wedge (empAge \leq depAge)} dependent)$

The above query evaluates to the set of **empId**s of employees whose age is greater than that of

A. some dependent.
B. all dependents.
C. some of his/her dependents.
D. all of his/her dependents.

gate2014-3 databases relational-algebra normal

3.14.21 Relational Algebra: GATE2015-1-7

https://gateoverflow.in/8094

SELECT operation in SQL is equivalent to

A. The selection operation in relational algebra
B. The selection operation in relational algebra, except that SELECT in SQL retains duplicates
C. The projection operation in relational algebra
D. The projection operation in relational algebra, except that SELECT in SQL retains duplicates

gate2015-1 databases sql relational-algebra easy

3.14.22 Relational Algebra: GATE2017-1-46

https://gateoverflow.in/118329

Consider a database that has the relation schema CR(StudentName, CourseName). An instance of the schema CR is as given below.

StudentName	CourseName
SA	CA
SA	CB
SA	CC
SB	CB
SB	CC
SC	CA
SC	CB
SC	CC
SD	CA
SD	CB
SD	CC
SD	CD
SE	CD
SE	CA
SE	CB
SF	CA
SF	CB
SF	CC

The following query is made on the database.

- $T1 \leftarrow \pi_{CourseName}\left(\sigma_{StudentName=SA}\left(CR\right)\right)$
- $T2 \leftarrow CR \div T1$

The number of rows in $T2$ is _____ .

gate2017-1 databases relational-algebra normal numerical-answers

3.14.23 Relational Algebra: GATE2018-41

https://gateoverflow.in/204115

Consider the relations $r(A, B)$ and $s(B, C)$, where $s.B$ is a primary key and $r.B$ is a foreign key referencing $s.B$. Consider the query

$Q : r \bowtie (\sigma_{B<5}(s))$

Let LOJ denote the natural left outer-join operation. Assume that r and s contain no null values.

Which of the following is NOT equivalent to Q?

A. $\sigma_{B<5}(r \bowtie s)$
B. $\sigma_{B<5}(r\ LOJ\ s)$
C. $r\ LOJ\ (\sigma_{B<5}(s))$
D. $\sigma_{B<5}(r)\ LOJ\ s$

gate2018 databases relational-algebra normal

3.14.24 Relational Algebra: GATE2019-55

https://gateoverflow.in/302793

Consider the following relations $P(X, Y, Z), Q(X, Y, T)$ and $R(Y, V)$.

Table: P

X	Y	Z
X1	Y1	Z1
X1	Y1	Z2
X2	Y2	Z2
X2	Y4	Z4

Table: Q

X	Y	T
X2	Y1	2
X1	Y2	5
X1	Y1	6
X3	Y3	1

Table: R

Y	V
Y1	V1
Y3	V2
Y2	V3
Y2	V2

How many tuples will be returned by the following relational algebra query?

$\Pi_x(\sigma_{(P.Y=R.Y \wedge R.V=V2))}(P \times R)) - \Pi_x(\sigma_{(Q.Y=R.Y \wedge Q.T>2))}(Q \times R))$

Answer: _____

gate2019 numerical-answers databases relational-algebra

3.14.25 Relational Algebra: TIFR2010-B-33

https://gateoverflow.in/19246

In a relational database there are three relations:

- Customers = C (C Name)
- Shops = S (S Name)
- Buys = B (C Name, S Name)

Then the Relational Algebra expression (Π is the projection operator).

$$C - \Pi_{CName}((C \times S) - B)$$

returns the names of

A. Customers who buy from at least one shop.
B. Customers who buy from at least two shops.
C. Customers who buy from all shops.
D. Customers who do not buy buy anything at all.
E. None of the above.

tifr2010 databases relational-algebra

3.14.26 Relational Algebra: TIFR2013-B-19

https://gateoverflow.in/25872

In a relational database there are three relations:

- Customers = C(CName),
- Shops = S(SName),
- Buys = B(CName, SName).

Which of the following relational algebra expressions returns the names of shops that have no customers at all? [Here Π is the projection operator.]

A. $\Pi_{SName} B$
B. $S - B$
C. $S - \Pi_{SName} B$
D. $S - \Pi_{SName}((C \times S) - B)$
E. None of the above

tifr2013 databases relational-algebra

3.15 Relational Calculus (14)

3.15.1 Relational Calculus: GATE1993-23

https://gateoverflow.in/2320

The following relations are used to store data about students, courses, enrollment of students in courses and teachers of courses. Attributes for primary key in each relation are marked by '*'.

```
Students (rollno*, sname, saddr)
courses (cno*, cname)
```

```
enroll(rollno*, cno*, grade)
teach(tno*, tname, cao*)
```

(cno is course number cname is course name, tno is teacher number, tname is teacher name, sname is student name, etc.)

Write a SQL query for retrieving roll number and name of students who got A grade in at least one course taught by teacher names Ramesh for the above relational database.

gate1993 databases sql relational-calculus normal

3.15.2 Relational Calculus: GATE1993-24

https://gateoverflow.in/203351

The following relations are used to store data about students, courses, enrollment of students in courses and teachers of courses. Attributes for primary key in each relation are marked by '*'.

```
Students (rollno*, sname, saddr)
courses (cno*, cname)
enroll(rollno*, cno*, grade)
teach(tno*, tname, cao*)
```

(cno is course number cname is course name, tno is teacher number, tname is teacher name, sname is student name, etc.)

For the relational database given above, the following functional dependencies hold:

rollno $\rightarrow$ sname, sdaddr cno $\rightarrow$ cname

tno $\rightarrow$ tname rollno, cno $\rightarrow$ grade

a. Is the database in 3rd normal form (3NF)?
b. If yes, prove that it is in 3 NF. If not normalize, the relations so that they are in 3NF (without proving)?

gate1993 databases sql relational-calculus normal

3.15.3 Relational Calculus: GATE1998-2.19

https://gateoverflow.in/1692

Which of the following query transformations (i.e., replacing the l.h.s. expression by the r.h.s expression) is incorrect? R_1 and R_2 are relations, C_1 and C_2 are selection conditions and A_1 and A_2 are attributes of R_1.

A. $\sigma_{C_1}(\sigma_{C_2}(R_1)) \rightarrow \sigma_{C_2}(\sigma_{C_1}(R_1))$
B. $\sigma_{C_1}(\pi_{A_1}(R_1)) \rightarrow \pi_{A_1}(\sigma_{C_1}(R_1))$
C. $\sigma_{C_1}(R_1 \cup R_2) \rightarrow \sigma_{C_1}(R_1) \cup \sigma_{C_1}(R_2)$
D. $\pi_{A_1}(\sigma_{C_1}(R_1)) \rightarrow \sigma_{C_1}(\pi_{A_1}(R_1))$

gate1998 databases relational-calculus normal

3.15.4 Relational Calculus: GATE1999-1.19

https://gateoverflow.in/1472

The relational algebra expression equivalent to the following tuple calculus expression:
$\{t \mid t \in r \wedge (t[A] = 10 \wedge t[B] = 20)\}$ is

A. $\sigma_{(A=10 \vee B=20)}(r)$
B. $\sigma_{(A=10)}(r) \cup \sigma_{(B=20)}(r)$
C. $\sigma_{(A=10)}(r) \cap \sigma_{(B=20)}(r)$
D. $\sigma_{(A=10)}(r) - \sigma_{(B=20)}(r)$

gate1999 databases relational-calculus normal

3.15.5 Relational Calculus: GATE2001-2.24

https://gateoverflow.in/742

Which of the rational calculus expression is not safe?

A. $\{t \mid \exists u \in R_1(t[A] = u[A]) \wedge \neg \exists s \in R_2(t[A] = s[A])\}$
B. $\{t \mid \forall u \in R_1(u[A] =$ " x " $\Rightarrow \exists s \in R_2(t[A] = s[A] \wedge s[A] = u[A]))\}$
C. $\{t \mid \neg(t \in R_1)\}$
D. $\{t \mid \exists u \in R_1(t[A] = u[A]) \wedge \exists s \in R_2(t[A] = s[A])\}$

gate2001 relational-calculus normal databases

3.15.6 Relational Calculus: GATE2002-1.20

https://gateoverflow.in/825

With regards to the expressive power of the formal relational query languages, which of the following statements is true?

A. Relational algebra is more powerful than relational calculus
B. Relational algebra has the same power as relational calculus
C. Relational algebra has the same power as safe relational calculus
D. None of the above

gate2002 databases relational-calculus normal

3.15.7 Relational Calculus: GATE2004-13

https://gateoverflow.in/1010

Let $R_1\left(\underline{A}, B, C\right)$ and $R_2\left(\underline{D}, E\right)$ be two relation schema, where the primary keys are shown underlined, and let C be a foreign key in R_1 referring to R_2. Suppose there is no violation of the above referential integrity constraint in the corresponding relation instances r_1 and r_2. Which of the following relational algebra expressions would necessarily produce an empty relation?

A. $\Pi_D(r_2) - \Pi_C(r_1)$
B. $\Pi_C(r_1) - \Pi_D(r_2)$
C. $\Pi_D\left(r_1 \bowtie_{C \neq D} r_2\right)$
D. $\Pi_C\left(r_1 \bowtie_{C=D} r_2\right)$

gate2004 databases relational-calculus easy

3.15.8 Relational Calculus: GATE2006-IT-15

https://gateoverflow.in/3554

Which of the following relational query languages have the same expressive power?

I. Relational algebra
II. Tuple relational calculus restricted to safe expressions
III. Domain relational calculus restricted to safe expressions

A. II and III only B. I and II only C. I and III only D. I, II and III

gate2006-it databases relational-algebra relational-calculus easy

3.15.9 Relational Calculus: GATE2007-60

https://gateoverflow.in/1258

Consider the relation **employee**(name, sex, supervisorName) with *name* as the key, *supervisorName* gives the name of the supervisor of the employee under consideration. What does the following Tuple Relational Calculus query produce?

$\{e.name \mid employee(e) \wedge (\forall x)\left[\neg employee\left(x\right) \vee x.supervisorName \neq e.name \vee x.sex = \text{``}male\text{''}\right]\}$

A. Names of employees with a male supervisor.
B. Names of employees with no immediate male subordinates.
C. Names of employees with no immediate female subordinates.
D. Names of employees with a female supervisor.

gate2007 databases relational-calculus normal

3.15.10 Relational Calculus: GATE2007-IT-65

https://gateoverflow.in/3510

Consider a selection of the form $\sigma_{A \leq 100}(r)$, where r is a relation with 1000 tuples. Assume that the attribute values for A among the tuples are uniformly distributed in the interval $[0, 500]$. Which one of the following options is the best estimate of the number of tuples returned by the given selection query ?

A. 50 B. 100 C. 150 D. 200

gate2007-it databases relational-calculus probability normal

3.15.11 Relational Calculus: GATE2008-15

https://gateoverflow.in/413

Which of the following tuple relational calculus expression(s) is/are equivalent to $\forall t \in r\left(P\left(t\right)\right)$?

I. $\neg \exists t \in r\left(P\left(t\right)\right)$
II. $\exists t \notin r\left(P\left(t\right)\right)$
III. $\neg \exists t \in r\left(\neg P\left(t\right)\right)$

IV. $\exists t \notin r\left(\neg P\left(t\right)\right)$

A. I only B. II only C. III only D. III and IV only

gate2008 databases relational-calculus normal

3.15.12 Relational Calculus: GATE2008-IT-75

https://gateoverflow.in/3389

Student (school-id, sch-roll-no, sname, saddress)
School (school-id, sch-name, sch-address, sch-phone)
Enrolment(school-id sch-roll-no, erollno, examname)
ExamResult(erollno, examname, marks)

Consider the following tuple relational calculus query.
{t | ∃ E ∈ Enrolment t = E.school-id ∧
| {x | x ∈ Enrolment ∧ x.school-id = t ∧

(∃ B ∈ ExamResult B.erollno = x.erollno
∧ B.examname = x.examname ∧ B.marks > 35)}| /

| {x | x ∈ Enrolment ∧ x.school-id = t}| * 100 > 35}

If a student needs to score more than 35 marks to pass an exam, what does the query return?

A. The empty set
B. schools with more than 35% of its students enrolled in some exam or the other
C. schools with a pass percentage above 35% over all exams taken together
D. schools with a pass percentage above 35% over each exam

gate2008-it databases relational-calculus normal

3.15.13 Relational Calculus: GATE2009-45

https://gateoverflow.in/1331

Let R and S be relational schemes such that $R = \{a, b, c\}$ and $S = \{c\}$. Now consider the following queries on the database:

1. $\pi_{R-S}(r) - \pi_{R-S}\left(\pi_{R-S}(r) \times s - \pi_{R-S,S}(r)\right)$
2. $\{t \mid t \in \pi_{R-S}(r) \wedge \forall u \in s\left(\exists v \in r\left(u = v[S] \wedge t = v\left[R - S\right]\right)\right)\}$
3. $\{t \mid t \in \pi_{R-S}(r) \wedge \forall v \in r\left(\exists u \in s\left(u = v[S] \wedge t = v\left[R - S\right]\right)\right)\}$
4.
```
Select R.a,R.b
    From R,S
    Where R.c = S.c
```

Which of the above queries are equivalent?

A. 1 and 2 B. 1 and 3 C. 2 and 4 D. 3 and 4

gate2009 databases relational-calculus difficult

3.15.14 Relational Calculus: GATE2013-35

https://gateoverflow.in/1546

Consider the following relational schema.

- Students(rollno: integer, sname: string)
- Courses(courseno: integer, cname: string)
- Registration(rollno: integer, courseno: integer, percent: real)

Which of the following queries are equivalent to this query in English?

"Find the distinct names of all students who score more than 90% in the course numbered 107"

I.
```
SELECT DISTINCT S.sname FROM Students as S, Registration
    as R WHERE
    R.rollno=S.rollno AND R.courseno=107 AND R.percent >90
```

II. $\prod_{sname}(\sigma_{courseno=107 \wedge percent>90}(Registration \bowtie Students))$

III. $\{T \mid \exists S \in Students, \exists R \in Registration(S.rollno = R.rollno$
$\wedge R.courseno = 107 \wedge R.percent > 90 \wedge T.sname = S.sname)\}$

IV. $\{\langle S_N \rangle \mid \exists S_R \exists R_P(\langle S_R, S_N \rangle \in Students \wedge \langle S_R, 107, R_P \rangle \in Registration \wedge R_P > 90)\}$

A. I, II, III and IV
B. I, II and III only
C. I, II and IV only
D. II, III and IV only

gate2013 databases sql relational-calculus normal

3.16 Safe Query (1)

3.16.1 Safe Query: GATE2017-1-41

https://gateoverflow.in/118324

Consider a database that has the relation schemas EMP(EmpId, EmpName, DeptId), and DEPT(DeptName, DeptId). Note that the DeptId can be permitted to be NULL in the relation EMP. Consider the following queries on the database expressed in tuple relational calculus.

I. $\{t \mid \exists u \in \text{EMP}(t[\text{EmpName}] = u[\text{EmpName}] \wedge \forall v \in \text{DEPT}(t[\text{DeptId}] \neq v[\text{DeptId}]))\}$
II. $\{t \mid \exists u \in \text{EMP}(t[\text{EmpName}] = u[\text{EmpName}] \wedge \exists v \in \text{DEPT}(t[\text{DeptId}] \neq v[\text{DeptId}]))\}$
III. $\{t \mid \exists u \in \text{EMP}(t[\text{EmpName}] = u[\text{EmpName}] \wedge \exists v \in \text{DEPT}(t[\text{DeptId}] = v[\text{DeptId}]))\}$

Which of the above queries are safe?

A. I and II only B. I and III only C. II and III only D. I, II and III

gate2017-1 databases relational-calculus safe-query normal

3.17 Sql (48)

3.17.1 Sql: GATE1988-12iii

https://gateoverflow.in/94625

Describe the relational algebraic expression giving the relation returned by the following SQL query.

```
Select      SNAME
from        S
Where       SNOin
            (select     SNO
            from        SP
            where       PNOin
                        (select     PNO
                        from        P
                        Where       COLOUR='BLUE'))
```

gate1988 normal descriptive databases sql

3.17.2 Sql: GATE1988-12iv

https://gateoverflow.in/94626

```
Select      SNAME
from        S
Where       SNOin
            (select     SNO
            from        SP
            where       PNOin
                        (select     PNO
                        from        P
                        Where       COLOUR='BLUE'))
```

What relations are being used in the above SQL query? Given at least two attributes of each of these relations.

gate1988 normal descriptive databases sql

3.17.3 Sql: GATE1990-10-a

https://gateoverflow.in/85686

Consider the following relational database:

- employees (eno, ename, address, basic-salary)
- projects (pno, pname, nos-of-staffs-allotted)
- working (pno, eno, pjob)

The queries regarding data in the above database are formulated below in SQL. Describe in ENGLISH sentences the two queries that have been posted:

i.
```
SELECT ename
FROM employees
WHERE eno IN
    (SELECT eno
```

```
    FROM working
    GROUP BY eno
    HAVING COUNT(*)=
        (SELECT COUNT(*)
        FROM projects))
```

ii.
```
SELECT pname
FROM projects
WHERE pno IN
    (SELECT pno
    FROM projects
    MINUS
    SELECT DISTINCT pno
    FROM working);
```

gate1990 descriptive databases sql

3.17.4 Sql: GATE1991-12,b

https://gateoverflow.in/42998

Suppose a database consist of the following relations:

```
SUPPLIER (SCODE,SNAME,CITY).

PART (PCODE,PNAME,PDESC,CITY).

PROJECTS (PRCODE,PRNAME,PRCITY).

SPPR (SCODE,PCODE,PRCODE,QTY).
```

Write algebraic solution to the following :
Get SCODE values for suppliers who supply to both projects PR1 and PR2.
Get PRCODE values for projects supplied by at least one supplier not in the same city.

sql gate1991 normal databases

3.17.5 Sql: GATE1991-12-a

https://gateoverflow.in/539

Suppose a database consist of the following relations:

```
SUPPLIER (SCODE,SNAME,CITY).
PART (PCODE,PNAME,PDESC,CITY).
PROJECTS (PRCODE,PRNAME,PRCITY).
SPPR (SCODE,PCODE,PRCODE,QTY).
```

Write SQL programs corresponding to the following queries:

i. Print PCODE values for parts supplied to any project in DEHLI by a supplier in DELHI.
ii. Print all triples <CITY, PCODE, CITY> such that a supplier in first city supplies the specified part to a project in the second city, but do not print the triples in which the two CITY values are same.

gate1991 databases sql normal

3.17.6 Sql: GATE1997-76-b

https://gateoverflow.in/203570

Consider the following relational database schema:

- EMP (eno name, age)
- PROJ (pno name)
- INVOLVED (eno, pno)

EMP contains information about employees. PROJ about projects and involved about which employees involved in which projects. The underlined attributes are the primary keys for the respective relations.

State in English (in not more than 15 words)

What the following relational algebra expressions are designed to determine

i. $\Pi_{eno}(INVOLVED) - \Pi_{eno}((\Pi_{eno}(INVOLVED) \times \Pi_{pno}(PROJ) - INVOLVED)$
ii. $\Pi_{age}(EMP) - \Pi_{age}(\sigma_{E.age<Emp.age}((\rho E(EMP) \times EMP))$

(Note: $\rho E(EMP)$ conceptually makes a copy of EMP and names it E (ρ is called the rename operator))

gate1997 databases sql

3.17.7 Sql: GATE1998-7-a

https://gateoverflow.in/1721

Suppose we have a database consisting of the following three relations.

- FREQUENTS (student, parlor) giving the parlors each student visits.
- SERVES (parlor, ice-cream) indicating what kind of ice-creams each parlor serves.
- LIKES (student, ice-cream) indicating what ice-creams each student likes.

(Assume that each student likes at least one ice-cream and frequents at least one parlor)

Express the following in SQL:

Print the students that frequent at least one parlor that serves some ice-cream that they like.

gate1998 databases sql descriptive

3.17.8 Sql: GATE1999-2.25

https://gateoverflow.in/1502

Which of the following is/are correct?

A. An SQL query automatically eliminates duplicates
B. An SQL query will not work if there are no indexes on the relations
C. SQL permits attribute names to be repeated in the same relation
D. None of the above

gate1999 databases sql easy

3.17.9 Sql: GATE1999-22-a

https://gateoverflow.in/1521

Consider the set of relations

- EMP (Employee-no. Dept-no, Employee-name, Salary)
- DEPT (Dept-no. Dept-name, Location)

Write an SQL query to:

a)Find all employees names who work in departments located at ‘Calcutta’ and whose salary is greater than Rs.50,000.

b)Calculate, for each department number, the number of employees with a salary greater than Rs. 1,00,000.

gate1999 databases sql easy

3.17.10 Sql: GATE1999-22-b

https://gateoverflow.in/203572

Consider the set of relations

- EMP (Employee-no. Dept-no, Employee-name, Salary)
- DEPT (Dept-no. Dept-name, Location)

Write an SQL query to:

Calculate, for each department number, the number of employees with a salary greater than Rs. 1,00,000

gate1999 databases sql easy

3.17.11 Sql: GATE2000-2.25

https://gateoverflow.in/672

Given relations r(w, x) and s(y, z) the result of

```
select distinct w, x
from r, s
```

is guaranteed to be same as r, provided.

A. r has no duplicates and s is non-empty
B. r and s have no duplicates
C. s has no duplicates and r is non-empty
D. r and s have the same number of tuples

gate2000 databases sql

3.17.12 Sql: GATE2000-2.26

https://gateoverflow.in/673

In SQL, relations can contain null values, and comparisons with null values are treated as unknown. Suppose all comparisons with a null value are treated as false. Which of the following pairs is not equivalent?

A. $x = 5 \quad not(not(x = 5))$
B. $x = 5 \quad x > 4$ and $x < 6$, where x is an integer
C. $x \neq 5 \quad not(x = 5)$
D. none of the above

gate2000 databases sql normal

3.17.13 Sql: GATE2000-22

https://gateoverflow.in/693

Consider a bank database with only one relation

transaction (transno, acctno, date, amount)

The amount attribute value is positive for deposits and negative for withdrawals.

a. Define an SQL view TP containing the information
(acctno,T1.date,T2.amount)
for every pair of transaction T1,T2 and such that T1 and T2 are transaction on the same account and the date of T2 is $\leq$ the date of T1.
b. Using only the above view TP, write a query to find for each account the minimum balance it ever reached (not including the 0 balance when the account is created). Assume there is at most one transaction per day on each account and each account has at least one transaction since it was created. To simplify your query, break it up into 2 steps by defining an intermediate view V.

gate2000 databases sql normal descriptive

3.17.14 Sql: GATE2001-2.25

https://gateoverflow.in/743

Consider a relation geq which represents "greater than or equal to", that is, $(x, y) \in$ geq only if $y \geq x$.

```
create table geq
(
    ib integer not null,
    ub integer not null,
    primary key ib,
    foreign key (ub) references geq on delete cascade
);
```

Which of the following is possible if tuple (x,y) is deleted?

A. A tuple (z,w) with z > y is deleted
B. A tuple (z,w) with z > x is deleted
C. A tuple (z,w) with w < x is deleted
D. The deletion of (x,y) is prohibited

gate2001 databases sql normal

3.17.15 Sql: GATE2001-21-a

https://gateoverflow.in/762

Consider a relation examinee (regno, name, score), where regno is the primary key to score is a real number.

Write a relational algebra using $(\Pi, \sigma, \rho, \times)$ to find the list of names which appear more than once in examinee.

gate2001 databases sql normal descriptive

3.17.16 Sql: GATE2001-21-b

https://gateoverflow.in/203574

Consider a relation examinee (regno, name, score), where regno is the primary key to score is a real number.

Write an SQL query to list the *regno* of examinees who have a score greater than the average score.

gate2001 databases sql normal descriptive

3.17.17 Sql: GATE2001-21-c

https://gateoverflow.in/203573

Consider a relation $\text{examinee}\,(\text{regno}, \text{name}, \text{score})$, where regno is the primary key to score is a real number.

Suppose the relation $\text{appears}\,(\text{regno}, \text{centr\_code})$ specifies the center where an examinee appears. Write an SQL query to list the centr_code having an examinee of score greater than 80.

gate2001 databases sql normal descriptive

3.17.18 Sql: GATE2003-86

https://gateoverflow.in/969

Consider the set of relations shown below and the SQL query that follows.

Students: (Roll_number, Name, Date_of_birth)

Courses: (Course_number, Course_name, Instructor)

Grades: (Roll_number, Course_number, Grade)

```
Select distinct Name
from Students, Courses, Grades
where Students.Roll_number=Grades.Roll_number
 and Courses.Instructor = 'Korth'
 and Courses.Course_number = Grades.Course_number
 and Grades.Grade = 'A'
```

Which of the following sets is computed by the above query?

A. Names of students who have got an A grade in all courses taught by Korth
B. Names of students who have got an A grade in all courses
C. Names of students who have got an A grade in at least one of the courses taught by Korth
D. None of the above

gate2003 databases sql easy

3.17.19 Sql: GATE2004-53

https://gateoverflow.in/1049

The employee information in a company is stored in the relation

- Employee (name, sex, salary, deptName)

Consider the following SQL query

```
Select deptName
    From Employee
    Where sex = 'M'
    Group by deptName
    Having avg(salary) >
        (select avg (salary) from Employee)
```

It returns the names of the department in which

A. the average salary is more than the average salary in the company
B. the average salary of male employees is more than the average salary of all male employees in the company
C. the average salary of male employees is more than the average salary of employees in same the department
D. the average salary of male employees is more than the average salary in the company

gate2004 databases sql normal

3.17.20 Sql: GATE2004-IT-74

https://gateoverflow.in/3718

A relational database contains two tables student and department in which student table has columns roll_no, name and dept_id and department table has columns dept_id and dept_name. The following insert statements were executed successfully to populate the empty tables:

```
Insert into department values (1, 'Mathematics')
Insert into department values (2, 'Physics')
Insert into student values (1, 'Navin', 1)
Insert into student values (2, 'Mukesh', 2)
Insert into student values (3, 'Gita', 1)
```

How many rows and columns will be retrieved by the following SQL statement?

```
Select * from student, department
```

A. 0 row and 4 columns
B. 3 rows and 4 columns

C. 3 rows and 5 columns
D. 6 rows and 5 columns

gate2004-it databases sql normal

3.17.21 Sql: GATE2004-IT-76

https://gateoverflow.in/3720

A table T1 in a relational database has the following rows and columns:

Roll no.	Marks
1	10
2	20
3	30
4	NULL

The following sequence of SQL statements was successfully executed on table T1.

```
Update T1 set marks = marks + 5
Select avg(marks) from T1
```

What is the output of the select statement?

A. 18.75 B. 20 C. 25 D. Null

gate2004-it databases sql normal

3.17.22 Sql: GATE2004-IT-78

https://gateoverflow.in/3722

Consider two tables in a relational database with columns and rows as follows:

Table: Student

Roll_no	Name	Dept_id
1	ABC	1
2	DEF	1
3	GHI	2
4	JKL	3

Table: Department

Dept_id	Dept_name
1	A
2	B
3	C

Roll_no is the primary key of the Student table, Dept_id is the primary key of the Department table and Student.Dept_id is a foreign key from Department.Dept_id
What will happen if we try to execute the following two SQL statements?

i. update Student set Dept_id = Null where Roll_on = 1
ii. update Department set Dept_id = Null where Dept_id = 1

A. Both i and ii will fail
B. i will fail but ii will succeed
C. i will succeed but ii will fail
D. Both i and ii will succeed

gate2004-it databases sql normal

3.17.23 Sql: GATE2005-77, ISRO2016-55

https://gateoverflow.in/1400

The relation **book** (title,price) contains the titles and prices of different books. Assuming that no two books have the same price, what does the following SQL query list?

```
select title
from book as B
where (select count(*)
    from book as T
    where T.price>B.price) < 5
```

A. Titles of the four most expensive books
B. Title of the fifth most inexpensive book
C. Title of the fifth most expensive book
D. Titles of the five most expensive books

gate2005 databases sql easy isro2016

3.17.24 Sql: GATE2005-IT-69

https://gateoverflow.in/3832

In an inventory management system implemented at a trading corporation, there are several tables designed to hold all the information. Amongst these, the following two tables hold information on which items are supplied by which suppliers, and which warehouse keeps which items along with the stock-level of these items.

Supply = (supplierid, itemcode)
Inventory = (itemcode, warehouse, stocklevel)

For a specific information required by the management, following SQL query has been written

```
Select distinct STMP.supplierid
From Supply as STMP
Where not unique (Select ITMP.supplierid
                  From Inventory, Supply as ITMP
                  Where STMP.supplierid = ITMP.supplierid
                  And ITMP.itemcode = Inventory.itemcode
                  And Inventory.warehouse = 'Nagpur');
```

For the warehouse at Nagpur, this query will find all suppliers who

A. do not supply any item
B. supply exactly one item
C. supply one or more items
D. supply two or more items

gate2005-it databases sql normal

3.17.25 Sql: GATE2006-67

https://gateoverflow.in/1845

Consider the relation account (customer, balance) where the customer is a primary key and there are no null values. We would like to rank customers according to decreasing balance. The customer with the largest balance gets rank 1. Ties are not broke but ranks are skipped: if exactly two customers have the largest balance they each get rank 1 and rank 2 is not assigned.

Query 1 :

select A.customer, count(B.customer)
from account A, account B
where A.balance <=B.balance
group by A.customer

Query 2 :

select A.customer, 1+count(B.customer)
from account A, account B
where A.balance < B.balance
group by A.customer

Consider these statements about Query 1 and Query 2.

1. Query 1 will produce the same row set as Query 2 for some but not all databases.
2. Both Query 1 and Query 2 are a correct implementation of the specification
3. Query1 is a correct implementation of the specification but Query 2 is not
4. Neither Query 1 nor Query 2 is a correct implementation of the specification
5. Assigning rank with a pure relational query takes less time than scanning in decreasing balance order assigning ranks using ODBC.

Which two of the above statements are correct?

A. 2 and 5 B. 1 and 3 C. 1 and 4 D. 3 and 5

gate2006 databases sql normal

3.17.26 Sql: GATE2006-68

https://gateoverflow.in/1846

Consider the relation enrolled (student, course) in which (student, course) is the primary key, and the relation paid (student, amount) where student is the primary key. Assume no null values and no foreign keys or integrity constraints. Given the following four queries:

Query1:

```
select student from enrolled where student in (select student from paid)
```

Query2:

```
select student from paid where student in (select student from enrolled)
```

Query3:

```
select E.student from enrolled E, paid P where E.student = P.student
```

Query4:

```
select student from paid where exists
    (select * from enrolled where enrolled.student = paid.student)
```

Which one of the following statements is correct?

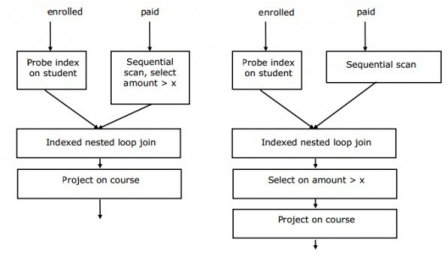
A. All queries return identical row sets for any database
B. Query2 and Query4 return identical row sets for all databases but there exist databases for which Query1 and Query2 return different row sets
C. There exist databases for which Query3 returns strictly fewer rows than Query2
D. There exist databases for which Query4 will encounter an integrity violation at runtime

gate2006 databases sql normal

3.17.27 Sql: GATE2006-69

https://gateoverflow.in/1847

Consider the relation enrolled (student, course) in which (student, course) is the primary key, and the relation paid (student, amount) where student is the primary key. Assume no null values and no foreign keys or integrity constraints. Assume that amounts 6000, 7000, 8000, 9000 and 10000 were each paid by 20% of the students. Consider these query plans (Plan 1 on left, Plan 2 on right) to *"list all courses taken by students who have paid more than x"*



A disk seek takes 4ms, disk data transfer bandwidth is 300 MB/s and checking a tuple to see if amount is greater than x takes 10µs. Which of the following statements is correct?

A. Plan 1 and Plan 2 will not output identical row sets for all databases
B. A course may be listed more than once in the output of Plan 1 for some databases
C. For x = 5000, Plan 1 executes faster than Plan 2 for all databases
D. For x = 9000, Plan I executes slower than Plan 2 for all databases

gate2006 databases sql normal

3.17.28 Sql: GATE2006-IT-84

https://gateoverflow.in/3640

Consider a database with three relation instances shown below. The primary keys for the Drivers and Cars relation are *did* and *cid* respectively and the records are stored in ascending order of these primary keys as given in the tables. No indexing is available in the database.

D: Drivers relation

did	dname	rating	age
22	Karthikeyan	7	25
29	Salman	1	33
31	Boris	8	55
32	Amoldt	8	25
58	Schumacher	10	35
64	Sachin	7	35
71	Senna	10	16
74	Sachin	9	35
85	Rahul	3	25
95	Ralph	3	53

R: Reserves relation

did	Cid	day
22	101	10 / 10 / 06
22	102	10 / 10 / 06
22	103	08 / 10 / 06
22	104	07 / 10 / 06
31	102	10 / 11 / 16
31	103	06 / 11 / 16
31	104	12 / 11 / 16
64	101	05 / 09 / 06
64	102	08 / 09 / 06
74	103	08 / 09 / 06

C: Cars relation

Cid	Cname	colour
101	Renault	blue
102	Renault	red
103	Ferrari	green
104	Jaguar	red

What is the output of the following SQL query?

```
select D.dname
from Drivers D
where D.did in  (
                        select R.did
                        from Cars C, Reserves R
                        where R.cid = C.cid and C.colour = 'red'
                        intersect
                        select R.did
                        from Cars C, Reserves R
                        where R.cid  = C.cid and C.colour = 'green'
                         )
```

A. Karthikeyan, Boris
B. Sachin, Salman
C. Karthikeyan, Boris, Sachin
D. Schumacher, Senna

gate2006-it databases sql normal

3.17.29 Sql: GATE2006-IT-85

https://gateoverflow.in/3641

Consider a database with three relation instances shown below. The primary keys for the Drivers and Cars relation are *did* and *cid* respectively and the records are stored in ascending order of these primary keys as given in the tables. No indexing is available in the database.

D: Drivers relation

did	dname	rating	age
22	Karthikeyan	7	25
29	Salman	1	33
31	Boris	8	55
32	Amoldt	8	25
58	Schumacher	10	35
64	Sachin	7	35
71	Senna	10	16
74	Sachin	9	35
85	Rahul	3	25
95	Ralph	3	53

R: Reserves relation

did	Cid	day
22	101	$10-10-06$
22	102	$10-10-06$
22	103	$08-10-06$
22	104	$07-10-06$
31	102	$10-11-16$
31	103	$06-11-16$
31	104	$12-11-16$
64	101	$05-09-06$
64	102	$08-09-06$
74	103	$08-09-06$

C: Cars relation

Cid	Cname	colour
101	Renault	blue
102	Renault	red
103	Ferrari	green
104	Jaguar	red

```
select D.dname
from Drivers D
where D.did in  (
                        select R.did
                        from Cars C, Reserves R
                        where R.cid = C.cid and C.colour = 'red'
                        intersect
                        select R.did
                        from Cars C, Reserves R
                        where R.cid  = C.cid and C.colour = 'green'
                         )
```

Let n be the number of comparisons performed when the above SQL query is optimally executed. If linear search is used to locate a tuple in a relation using primary key, then n lies in the range:

A. $36-40$ B. $44-48$ C. $60-64$ D. $100-104$

gate2006-it databases sql normal

3.17.30 Sql: GATE2007-61

https://gateoverflow.in/1259

Consider the table **employee**(empId, name, department, salary) and the two queries Q_1, Q_2 below. Assuming that department 5 has more than one employee, and we want to find the employees who get higher salary than anyone in the department 5, which one of the statements is **TRUE** for any arbitrary employee table?

Q_1 :
```
Select e.empId
From employee e
Where not exists
    (Select * From employee s Where s.department = "5" and s.salary >= e.salary)
```

Q_2 :
```
Select e.empId
From employee e
Where e.salary > Any
    (Select distinct salary From employee s Where s.department = "5")
```

A. Q_1 is the correct query
B. Q_2 is the correct query
C. Both Q_1 and Q_2 produce the same answer
D. Neither Q_1 nor Q_2 is the correct query

gate2007 databases sql normal verbal-ability

3.17.31 Sql: GATE2008-IT-74

https://gateoverflow.in/3388

Student (school-id, sch-roll-no, sname, saddress)
School (school-id, sch-name, sch-address, sch-phone)
Enrolment(school-id sch-roll-no, erollno, examname)
ExamResult(erollno, examname, marks)

What does the following SQL query output?

```
SELECT sch-name, COUNT (*)
FROM School C, Enrolment E, ExamResult R
WHERE E.school-id = C.school-id
AND
E.examname = R.examname AND E.erollno = R.erollno
AND
R.marks = 100 AND S.school-id IN (SELECT school-id
                                  FROM student
                                  GROUP BY school-id
                                   HAVING COUNT (*) > 200)
GROUP By school-id
```

A. for each school with more than 200 students appearing in exams, the name of the school and the number of 100s scored by its students
B. for each school with more than 200 students in it, the name of the school and the number of 100s scored by its students
C. for each school with more than 200 students in it, the name of the school and the number of its students scoring 100 in at least one exam
D. nothing; the query has a syntax error

gate2008-it databases sql normal

3.17.32 Sql: GATE2009-55

https://gateoverflow.in/1339

Consider the following relational schema:

Suppliers(sid:integer, sname:string, city:string, street:string)

Parts(pid:integer, pname:string, color:string)

Catalog(sid:integer, pid:integer, cost:real)

Consider the following relational query on the above database:

```
SELECT S.sname
FROM    Suppliers S
WHERE S.sid NOT IN (SELECT C.sid
                    FROM Catalog C
                    WHERE C.pid NOT IN (SELECT P.pid
                                        FROM Parts P
                                        WHERE P.color<>'blue'))
```

Assume that relations corresponding to the above schema are not empty. Which one of the following is the correct interpretation of the above query?

A. Find the names of all suppliers who have supplied a non-blue part.
B. Find the names of all suppliers who have not supplied a non-blue part.
C. Find the names of all suppliers who have supplied only non-blue part.
D. Find the names of all suppliers who have not supplied only blue parts.

gate2009 databases sql normal

3.17.33 Sql: GATE2010-19

https://gateoverflow.in/2194

A relational schema for a train reservation database is given below.

- **passenger(pid, pname, age)**
- **reservation(pid, class, tid)**

Passenger

pid	pname	Age
0	Sachine	65
1	Rahul	66
2	Sourav	67
3	Anil	69

Reservation

pid	class	tid
0	AC	8200
1	AC	8201
2	SC	8201
5	AC	8203
1	SC	8204
3	AC	8202

What **pids** are returned by the following SQL query for the above instance of the tables?

```
SELECT pid
FROM Reservation
WHERE class='AC' AND
    EXISTS (SELECT *
            FROM Passenger
            WHERE age>65 AND
            Passenger.pid=Reservation.pid)
```

A. $1, 0$ B. $1, 2$ C. $1, 3$ D. $1, 5$

gate2010 databases sql normal

3.17.34 Sql: GATE2011-32 https://gateoverflow.in/2134

Consider a database table T containing two columns X and Y each of type $\mathrm{integer}$. After the creation of the table, one record $(\mathrm{X}{=}1,\ \mathrm{Y}{=}1)$ is inserted in the table.

Let MX and MY denote the respective maximum values of X and Y among all records in the table at any point in time. Using MX and MY, new records are inserted in the table 128 times with X and Y values being $\mathrm{MX}{+}1$, $2{*}\mathrm{MY}{+}1$ respectively. It may be noted that each time after the insertion, values of MX and MY change.

What will be the output of the following SQL query after the steps mentioned above are carried out?

```
SELECT Y FROM T WHERE X=7;
```

A. 127 B. 255 C. 129 D. 257

gate2011 databases sql normal

3.17.35 Sql: GATE2011-46 https://gateoverflow.in/2148

Database table by name $\mathrm{Loan\_Records}$ is given below.

Borrower	Bank_Manager	Loan_Amount
Ramesh	Sunderajan	10000.00
Suresh	Ramgopal	5000.00
Mahesh	Sunderajan	7000.00

What is the output of the following SQL query?

```
SELECT count(*)
FROM (
    SELECT Borrower, Bank_Manager FROM Loan_Records) AS S
    NATURAL JOIN
    (SELECT Bank_Manager, Loan_Amount FROM Loan_Records) AS T
);
```

A. 3 B. 9 C. 5 D. 6

gate2011 databases sql normal

3.17.36 Sql: GATE2012-15 https://gateoverflow.in/47

Which of the following statements are **TRUE** about an SQL query?

P : An SQL query can contain a HAVING clause even if it does not have a GROUP BY clause
Q : An SQL query can contain a HAVING clause only if it has a GROUP BY clause
R : All attributes used in the GROUP BY clause must appear in the SELECT clause
S : Not all attributes used in the GROUP BY clause need to appear in the SELECT clause

A. P and R B. P and S C. Q and R D. Q and S

gate2012 databases easy sql ambiguous

3.17.37 Sql: GATE2012-51

https://gateoverflow.in/43313

Consider the following relations A, B and C :

A

Id	Name	Age
12	Arun	60
15	Shreya	24
99	Rohit	11

B

Id	Name	Age
15	Shreya	24
25	Hari	40
98	Rohit	20
99	Rohit	11

C

Id	Phone	Area
10	2200	02
99	2100	01

How many tuples does the result of the following SQL query contain?

```
SELECT A.Id
FROM A
WHERE A.Age > ALL (SELECT B.Age
                   FROM B
                   WHERE B.Name = 'Arun')
```

A. 4 B. 3 C. 0 D. 1

gate2012 databases sql normal

3.17.38 Sql: GATE2014-1-22

https://gateoverflow.in/1789

Given the following statements:

S1: A foreign key declaration can always be replaced by an equivalent check assertion in SQL.

S2: Given the table $R(a, b, c)$ where a and b together form the primary key, the following is a valid table definition.

```
CREATE TABLE S (
     a INTEGER,
     d INTEGER,
     e INTEGER,
     PRIMARY KEY (d),
     FOREIGN KEY (a) references R)
```

Which one of the following statements is **CORRECT**?

A. S1 is TRUE and S2 is FALSE
B. Both S1 and S2 are TRUE
C. S1 is FALSE and S2 is TRUE
D. Both S1 and S2 are FALSE

gate2014-1 databases normal sql

3.17.39 Sql: GATE2014-1-54

https://gateoverflow.in/1934

Given the following schema:

employees(emp-id, first-name, last-name, hire-date, dept-id, salary)

departments(dept-id, dept-name, manager-id, location-id)

You want to display the last names and hire dates of all latest hires in their respective departments in the location ID 1700. You issue the following query:

```
SQL>SELECT last-name, hire-date
    FROM employees
    WHERE (dept-id, hire-date) IN
    (SELECT dept-id, MAX(hire-date)
    FROM employees JOIN departments USING(dept-id)
    WHERE location-id =1700
    GROUP BY dept-id);
```

What is the outcome?

A. It executes but does not give the correct result
B. It executes and gives the correct result.
C. It generates an error because of pairwise comparison.
D. It generates an error because of the GROUP BY clause cannot be used with table joins in a sub-query.

gate2014-1 databases sql normal

3.17.40 Sql: GATE2014-2-54

https://gateoverflow.in/2021

SQL allows duplicate tuples in relations, and correspondingly defines the multiplicity of tuples in the result of joins. Which one of the following queries always gives the same answer as the nested query shown below:

```
select * from R where a in (select S.a from S)
```

A. select R.* from R, S where R.a=S.a
B. select distinct R.* from R,S where R.a=S.a
C. select R.* from R,(select distinct a from S) as S1 where R.a=S1.a
D. select R.* from R,S where R.a=S.a and is unique R

gate2014-2 databases sql normal

3.17.41 Sql: GATE2014-3-54

https://gateoverflow.in/2089

Consider the following relational schema:

employee (empId,empName,empDept)

customer (custId,custName,salesRepId,rating)

salesRepId is a foreign key referring to **empId** of the employee relation. Assume that each employee makes a sale to at least one customer. What does the following query return?

```
SELECT empName   FROM employee E
WHERE NOT EXISTS (SELECT custId
    FROM customer C
    WHERE C.salesRepId = E.empId
    AND C.rating <> 'GOOD');
```

A. Names of all the employees with at least one of their customers having a ‘GOOD’ rating.
B. Names of all the employees with at most one of their customers having a 'GOOD' rating.
C. Names of all the employees with none of their customers having a 'GOOD' rating.
D. Names of all the employees with all their customers having a 'GOOD' rating.

gate2014-3 databases sql easy

3.17.42 Sql: GATE2015-1-27

https://gateoverflow.in/8225

Consider the following relation:

Student

Roll_No	Student_Name
1	Raj
2	Rohit
3	Raj

Performance

Roll_No	Course	Marks
1	Math	80
1	English	70
2	Math	75
3	English	80
2	Physics	65
3	Math	80

Consider the following SQL query.

```
SELECT S.Student_Name, Sum(P. Marks)
FROM Student S, Performance P
WHERE S.Roll_No= P.Roll_No
GROUP BY S.STUDENT_Name
```

The numbers of rows that will be returned by the SQL query is_____.

gate2015-1 databases sql normal numerical-answers

3.17.43 Sql: GATE2015-3-3

https://gateoverflow.in/8396

Consider the following relation

Cinema(*theater, address, capacity*)

Which of the following options will be needed at the end of the SQL query

```
SELECT P1.address
FROM Cinema P1
```

such that it always finds the addresses of theaters with maximum capacity?

A.
```
WHERE P1.capacity >= All (select P2.capacity from Cinema P2)
```

B.
```
WHERE P1.capacity >= Any (select P2.capacity from Cinema P2)
```

C.
```
WHERE P1.capacity > All (select max(P2.capacity) from Cinema P2)
```

D.
```
WHERE P1.capacity > Any (select max(P2.capacity) from Cinema P2)
```

gate2015-3 databases sql normal

3.17.44 Sql: GATE2016-2-52

https://gateoverflow.in/39604

Consider the following database table named water_schemes:

Water_schemes

scheme_no	district_name	capacity
1	Ajmer	20
1	Bikaner	10
2	Bikaner	10
3	Bikaner	20
1	Churu	10
2	Churu	20
1	Dungargarh	10

The number of tuples returned by the following SQL query is _____.

```
with total (name, capacity) as
    select district_name, sum (capacity)
    from water_schemes
    group by district_name
with total_avg (capacity) as
   select avg (capacity)
   from total
select name
   from total, total_avg
   where total.capacity ≥ total_avg.capacity
```

gate2016-2 databases sql normal numerical-answers

3.17.45 Sql: GATE2017-1-23

https://gateoverflow.in/118303

Consider a database that has the relation schema EMP (EmpId, EmpName, and DeptName). An instance of the schema EMP and a SQL query on it are given below:

EMP

EmpId	EmpName	DeptName
1	XYA	AA
2	XYB	AA
3	XYC	AA
4	XYD	AA
5	XYE	AB
6	XYF	AB
7	XYG	AB
8	XYH	AC
9	XYI	AC
10	XYJ	AC
11	XYK	AD
12	XYL	AD
13	XYM	AE

```
SELECT AVG(EC.Num)
FROM EC
WHERE (DeptName, Num) IN
    (SELECT DeptName, COUNT(EmpId) AS
                        EC(DeptName, Num)
    FROM EMP
    GROUP BY DeptName)
```

The output of executing the SQL query is _____ .

gate2017-1 databases sql numerical-answers

3.17.46 Sql: GATE2017-2-46

https://gateoverflow.in/118391

Consider the following database table named top_scorer.

top_scorer

player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
G Muller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
T Muller	Germany	10
Rahn	Germany	10

Consider the following SQL query:

```
SELECT ta.player FROM top_scorer AS ta
WHERE ta.goals >ALL (SELECT tb.goals
    FROM top_scorer AS tb
    WHERE tb.country = 'Spain')
AND ta.goals > ANY (SELECT tc.goals
    FROM top_scorer AS tc
    WHERE tc.country='Germany')
```

The number of tuples returned by the above SQL query is _____

gate2017-2 databases sql numerical-answers

3.17.47 Sql: GATE2018-12

https://gateoverflow.in/204086

Consider the following two tables and four queries in SQL.

Book (isbn, bname), Stock(isbn, copies)

Query 1:

```
SELECT B.isbn, S.copies FROM Book B INNER JOIN Stock S ON B.isbn=S.isbn;
```

Query 2:

```
SELECT B.isbn, S.copies FROM Book B LEFT OUTER JOIN Stock S ON B.isbn=S.isbn;
```

Query 3:

```
SELECT B.isbn, S,copies FROM Book B RIGHT OUTER JOIN Stock S ON B.isbn=S.isbn
```

Query 4:

```
SELECT B.isbn, S.copies FROM Book B FULL OUTER JOIN Stock S ON B.isbn=S.isbn
```

Which one of the queries above is certain to have an output that is a superset of the outputs of the other three queries?

A. Query 1 B. Query 2 C. Query 3 D. Query 4

gate2018 databases sql easy

3.17.48 Sql: GATE2019-51

https://gateoverflow.in/302797

A relational database contains two tables Student and Performance as shown below:

Table: student

Roll_no	Student_name
1	Amit
2	Priya
3	Vinit
4	Rohan
5	Smita

Table: Performance

Roll_no	Subject_code	Marks
1	A	86
1	B	95
1	C	90
2	A	89
2	C	92
3	C	80

The primary key of the Student table is Roll_no. For the performance table, the columns Roll_no. and Subject_code together form the primary key. Consider the SQL query given below:

```
SELECT S.Student_name, sum(P.Marks) FROM Student S, Performance P WHERE P.Marks >84 GROUP BY S.Student_name;
```

The number of rows returned by the above SQL query is _____

gate2019 numerical-answers databases sql

3.18 Timestamp Ordering (1)

3.18.1 Timestamp Ordering: GATE2017-1-42

https://gateoverflow.in/118325

In a database system, unique timestamps are assigned to each transaction using Lamport's logical clock. Let $TS(T_1)$ and $TS(T_2)$ be the timestamps of transactions T_1 and T_2 respectively. Besides, T_1 holds a lock on the resource R, and T_2 has requested a conflicting lock on the same resource R. The following algorithm is used to prevent deadlocks in the database system assuming that a killed transaction is restarted with the same timestamp.

if $TS(T_2) < TS(T_1)$ then

T_1 is killed

else T_2 waits.

Assume any transaction that is not killed terminates eventually. Which of the following is TRUE about the database system that uses the above algorithm to prevent deadlocks?

A. The database system is both deadlock-free and starvation-free.
B. The database system is deadlock-free, but not starvation-free.
C. The database system is starvation-free, but not deadlock-free.
D. The database system is neither deadlock-free nor starvation-free.

gate2017-1 databases timestamp-ordering deadlock normal

3.19 Transaction And Concurrency (1)

3.19.1 Transaction And Concurrency: GATE2004-IT-21

https://gateoverflow.in/3662

Which level of locking provides the highest degree of concurrency in a relational database ?

A. Page
B. Table
C. Row
D. Page, table and row level locking allow the same degree of concurrency

gate2004-it databases normal transaction-and-concurrency

3.20 Transactions (25)

3.20.1 Transactions: GATE1999-2.6

https://gateoverflow.in/1484

For the schedule given below, which of the following is correct:

1	Read A	
2		Read B
3	Write A	
4		Read A
5		Write A
6		Write B
7	Read B	
8	Write B	

A. This schedule is serializable and can occur in a scheme using 2PL protocol
B. This schedule is serializable but cannot occur in a scheme using 2PL protocol
C. This schedule is not serializable but can occur in a scheme using 2PL protocol
D. This schedule is not serializable and cannot occur in a scheme using 2PL protocol

gate1999 databases transactions normal

3.20.2 Transactions: GATE2003-29, ISRO2009-73

https://gateoverflow.in/919

Which of the following scenarios may lead to an irrecoverable error in a database system?

A. A transaction writes a data item after it is read by an uncommitted transaction
B. A transaction reads a data item after it is read by an uncommitted transaction
C. A transaction reads a data item after it is written by a committed transaction
D. A transaction reads a data item after it is written by an uncommitted transaction

gate2003 databases transactions easy isro2009

3.20.3 Transactions: GATE2003-87

https://gateoverflow.in/970

Consider three data items $D1, D2$, and $D3$, and the following execution schedule of transactions $T1, T2$, and $T3$. In the diagram, $R(D)$ and $W(D)$ denote the actions reading and writing the data item D respectively.

time ↓

Which of the following statements is correct?

A. The schedule is serializable as $T2; T3; T1$
B. The schedule is serializable as $T2; T1; T3$
C. The schedule is serializable as $T3; T2; T1$
D. The schedule is not serializable

gate2003 databases transactions normal

3.20.4 Transactions: GATE2004-IT-77

https://gateoverflow.in/3721

Consider the following schedule S of transactions $T1$ and $T2$:

T1	T2
Read(A)	
A = A – 10	
	Read(A)
	Temp = 0.2*A
	Write(A)
	Read(B)
Write(A)	
Read(B)	
B = B + 10	
Write(B)	
	B = B + Temp
	Write(B)

Which of the following is TRUE about the schedule S ?

A. S is serializable only as $T1, T2$
B. S is serializable only as $T2, T1$
C. S is serializable both as $T1, T2$ and $T2, T1$
D. S is not serializable either as $T1, T2$ or as $T2, T1$

gate2004-it databases transactions normal

3.20.5 Transactions: GATE2005-IT-24

https://gateoverflow.in/3769

Amongst the ACID properties of a transaction, the 'Durability' property requires that the changes made to the database by a successful transaction persist

A. Except in case of an Operating System crash
B. Except in case of a Disk crash
C. Except in case of a power failure
D. Always, even if there is a failure of any kind

gate2005-it databases transactions easy

3.20.6 Transactions: GATE2005-IT-67

https://gateoverflow.in/3830

A company maintains records of sales made by its salespersons and pays them commission based on each individual's total sales made in a year. This data is maintained in a table with following schema:

salesinfo = (salespersonid, totalsales, commission)

In a certain year, due to better business results, the company decides to further reward its salespersons by enhancing the commission paid to them as per the following formula:

If commission ≤ 50000, enhance it by 2%
If $50000 <$ commission ≤ 100000, enhance it by 4%
If commission > 100000, enhance it by 6%

The IT staff has written three different SQL scripts to calculate enhancement for each slab, each of these scripts is to run as a separate transaction as follows:

T1
```
Update salesinfo
Set commission = commission * 1.02
Where commission < = 50000;
```

T2
```
Update salesinfo
Set commission = commission * 1.04
Where commission > 50000 and commission is < = 100000;
```

T3
```
Update salesinfo
Set commission = commission * 1.06
Where commission > 100000;
```

Which of the following options of running these transactions will update the commission of all salespersons correctly

A. Execute T1 followed by T2 followed by T3
B. Execute T2, followed by T3; T1 running concurrently throughout
C. Execute T3 followed by T2; T1 running concurrently throughout
D. Execute T3 followed by T2 followed by T1

gate2005-it databases transactions normal

3.20.7 Transactions: GATE2006-20, ISRO2015-17

https://gateoverflow.in/981

Consider the following log sequence of two transactions on a bank account, with initial balance 12000, that transfer 2000 to a mortgage payment and then apply a 5% interest.

1. T1 start
2. T1 B old $= 1200$ new $= 10000$
3. T1 M old $= 0$ new $= 2000$
4. T1 commit
5. T2 start
6. T2 B old $= 10000$ new $= 10500$
7. T2 commit

Suppose the database system crashes just before log record 7 is written. When the system is restarted, which one statement is true of the recovery procedure?

A. We must redo log record 6 to set B to 10500
B. We must undo log record 6 to set B to 10000 and then redo log records 2 and 3
C. We need not redo log records 2 and 3 because transaction T1 has committed
D. We can apply redo and undo operations in arbitrary order because they are idempotent

gate2006 databases transactions normal isro2015

3.20.8 Transactions: GATE2007-64

https://gateoverflow.in/1262

Consider the following schedules involving two transactions. Which one of the following statements is **TRUE**?

- $S_1 : r_1(X); r_1(Y); r_2(X); r_2(Y); w_2(Y); w_1(X)$
- $S_2 : r_1(X); r_2(X); r_2(Y); w_2(Y); r_1(Y); w_1(X)$

A. Both S_1 and S_2 are conflict serializable.
B. S_1 is conflict serializable and S_2 is not conflict serializable.
C. S_1 is not conflict serializable and S_2 is conflict serializable.
D. Both S_1 and S_2 are not conflict serializable.

gate2007 databases transactions normal

3.20.9 Transactions: GATE2007-IT-66

https://gateoverflow.in/3511

Consider the following two transactions : T_1 and T2.

T_1 : read (A);
read (B);
if A = 0 then B ← B + 1;
write (B);

T_2 : read (B);
read (A);
if B ≠ 0 then A ← A - 1;
write (A);

Which of the following schemes, using shared and exclusive locks, satisfy the requirements for strict two phase locking for the

above transactions?

A.

S1 :	S2 :
lock S(A);	lock S(B);
read (A);	read (B);
lock S(B);	lock S(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B ← B + 1;	then A ← A - 1;
write (B);	write (A);
commit;	commit;
unlock (A);	unlock (B);
unlock (B);	unlock (A);

B.

S1 :	S2 :
lock X(A);	lock X(B);
read (A);	read (B);
lock X(B);	lock X(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B ← B + 1;	then A ← A - 1;
write (B);	write (A);
unlock (A);	unlock (A);
commit;	commit;
unlock (B);	unlock (A);

C.

S1 :	S2 :
lock S(A);	lock S(B);
read (A);	read (B);
lock X(B);	lock X(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B ← B + 1;	then A ← A - 1;
write (B);	write (A);
unlock (A);	unlock (B);
commit;	commit;
unlock (B);	unlock (A);

D.

S1 :	S2 :
lock S(A);	lock S(B);
read (A);	read (B);
lock X(B);	lock X(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B ← B + 1;	then A ← A - 1;
write (B);	write (A);
unlock (A);	unlock (A);
unlock (B);	unlock (B);
commit;	commit;

gate2007-it databases transactions normal

3.20.10 Transactions: GATE2008-IT-63

https://gateoverflow.in/3374

Consider the following three schedules of transactions T1, T2 and T3. [Notation: In the following NYO represents the action Y (R for read, W for write) performed by transaction N on object O.]

$(S1)$	$2RA$	$2WA$	$3RC$	$2WB$	$3WA$	$3WC$	$1RA$	$1RB$	$1WA$	$1WB$
$(S2)$	$3RC$	$2RA$	$2WA$	$2WB$	$3WA$	$1RA$	$1RB$	$1WA$	$1WB$	$3WC$
$(S3)$	$2RA$	$3RC$	$3WA$	$2WA$	$2WB$	$3WC$	$1RA$	$1RB$	$1WA$	$1WB$

Which of the following statements is TRUE?

A. S1, S2 and S3 are all conflict equivalent to each other
B. No two of S1, S2 and S3 are conflict equivalent to each other
C. S2 is conflict equivalent to S3, but not to S1

D. S1 is conflict equivalent to S2, but not to S3

gate2008-it databases transactions normal

3.20.11 Transactions: GATE2009-43

https://gateoverflow.in/1329

Consider two transactions T_1 and T_2, and four schedules S_1, S_2, S_3, S_4, of T_1 and T_2 as given below:

$T_1 : R_1[x]W_1[x]W_1[y]$

$T_2 : R_2[x]R_2[y]W_2[y]$

$S_1 : R_1[x]R_2[x]R_2[y]W_1[x]W_1[y]W_2[y]$

$S_2 : R_1[x]R_2[x]R_2[y]W_1[x]W_2[y]W_1[y]$

$S_3 : R_1[x]W_1[x]R_2[x]W_1[y]R_2[y]W_2[y]$

$S_4 : R_2[x]R_2[y]R_1[x]W_1[x]W_1[y]W_2[y]$

Which of the above schedules are conflict-serializable?

A. S_1 and S_2 B. S_2 and S_3 C. S_3 only D. S_4 only

gate2009 databases transactions normal

3.20.12 Transactions: GATE2010-20

https://gateoverflow.in/2196

Which of the following concurrency control protocols ensure both conflict serializability and freedom from deadlock?

I. 2-phase locking
II. Time-stamp ordering

A. I only B. II only C. Both I and II D. Neither I nor II

gate2010 databases transactions normal

3.20.13 Transactions: GATE2010-42

https://gateoverflow.in/2343

Consider the following schedule for transactions $T1, T2$ and $T3$:

T1	T2	T3
Read(X)		
	Read(Y)	
		Read(Y)
	Write(Y)	
Write(X)		
		Write(X)
	Read(X)	
	Write(X)	

Which one of the schedules below is the correct serialization of the above?

A. $T1 \rightarrow T3 \rightarrow T2$ B. $T2 \rightarrow T1 \rightarrow T3$
C. $T2 \rightarrow T3 \rightarrow T1$ D. $T3 \rightarrow T1 \rightarrow T2$

gate2010 databases transactions normal

3.20.14 Transactions: GATE2012-27

https://gateoverflow.in/1612

Consider the following transactions with data items P and Q initialized to zero:

T_1	read (P); read (Q); if P = 0 then Q := Q + 1; write (Q)
T_2	read (Q); read (P); if Q = 0 then P := P + 1; write (P)

Any non-serial interleaving of **T1** and **T2** for concurrent execution leads to

A. a serializable schedule
B. a schedule that is not conflict serializable
C. a conflict serializable schedule
D. a schedule for which a precedence graph cannot be drawn

gate2012 databases transactions normal

3.20.15 Transactions: GATE2014-1-29

https://gateoverflow.in/1796

Consider the following four schedules due to three transactions (indicated by the subscript) using *read* and *write* on a data item x, denoted by $r(x)$ and $w(x)$ respectively. Which one of them is conflict serializable?

A. $r_1(x); r_2(x); w_1(x); r_3(x); w_2(x);$
B. $r_2(x); r_1(x); w_2(x); r_3(x); w_1(x);$
C. $r_3(x); r_2(x); r_1(x); w_2(x); w_1(x);$
D. $r_2(x); w_2(x); r_3(x); r_1(x); w_1(x);$

gate2014-1 databases transactions normal

3.20.16 Transactions: GATE2014-2-29

https://gateoverflow.in/1988

Consider the following schedule **S** of transactions $T1, T2, T3, T4$:

T1	T2	T3	T4
	Reads(X)		
		Writes(X)	
		Commit	
Writes(X)			
Commit			
	Writes(Y)		
	Reads(Z)		
	Commit		
			Reads(X)
			Reads(Y)
			Commit

Which one of the following statements is CORRECT?

A. **S** is conflict-serializable but not recoverable
B. **S** is not conflict-serializable but is recoverable
C. **S** is both conflict-serializable and recoverable
D. **S** is neither conflict-serializable not is it recoverable

gate2014-2 databases transactions normal

3.20.17 Transactions: GATE2014-3-29

https://gateoverflow.in/2063

Consider the transactions $T1, T2,$ and $T3$ and the schedules $S1$ and $S2$ given below.

$T1 : r1(X); r1(Z); w1(X); w1(Z)$

$T2 : r2(Y); r2(Z); w2(Z)$

$T3 : r3(Y); r3(X); w3(Y)$

$S1 : r1(X); r3(Y); r3(X); r2(Y); r2(Z); w3(Y); w2(Z); r1(Z); w1(X); w1(Z)$

$S2 : r1(X); r3(Y); r2(Y); r3(X); r1(Z); r2(Z); w3(Y); w1(X); w2(Z); w1(Z)$

Which one of the following statements about the schedules is **TRUE**?

A. Only $S1$ is conflict-serializable.
B. Only $S2$ is conflict-serializable.
C. Both $S1$ and $S2$ are conflict-serializable.
D. Neither $S1$ nor $S2$ is conflict-serializable.

gate2014-3 databases transactions normal

3.20.18 Transactions: GATE2015-2-1

https://gateoverflow.in/8047

Consider the following transaction involving two bank accounts x and y.

```
read(x); x:=x-50; write (x); read(y); y:=y+50; write(y)
```

The constraint that the sum of the accounts x and y should remain constant is that of

A. Atomicity B. Consistency C. Isolation D. Durability

gate2015-2 databases transactions easy

3.20.19 Transactions: GATE2015-2-46

https://gateoverflow.in/8246

Consider a simple checkpointing protocol and the following set of operations in the log.

(start, T4); (write, T4, y, 2, 3); (start, T1); (commit, T4); (write, T1, z, 5, 7);

(checkpoint);

(start, T2); (write, T2, x, 1, 9); (commit, T2); (start, T3); (write, T3, z, 7, 2);

If a crash happens now and the system tries to recover using both undo and redo operations, what are the contents of the undo list and the redo list?

A. Undo: T3, T1; Redo: T2
B. Undo: T3, T1; Redo: T2, T4
C. Undo: none; Redo: T2, T4, T3, T1
D. Undo: T3, T1, T4; Redo: T2

gate2015-2 databases transactions normal

3.20.20 Transactions: GATE2015-3-29

https://gateoverflow.in/8482

Consider the partial Schedule S involving two transactions $T1$ and $T2$. Only the *read* and the *write* operations have been shown. The *read* operation on data item P is denoted by *read(P)* and *write* operation on data item P is denoted by *write(P)*.

Schedule S

Time Instance	Transaction ID	
	T1	**T2**
1	read(A)	
2	write(A)	
3		read(C)
4		write(C)
5		read(B)
6		write(B)
7		read(A)
8		commit
9	read(B)	

Suppose that the transaction $T1$ fails immediately after time instance 9. Which of the following statements is correct?

A. $T2$ must be aborted and then both $T1$ and $T2$ must be re-started to ensure transaction atomicity
B. Schedule S is non-recoverable and cannot ensure transaction atomicity
C. Only $T2$ must be aborted and then re-started to ensure transaction atomicity
D. Schedule S is recoverable and can ensure transaction atomicity and nothing else needs to be done

gate2015-3 databases transactions normal

3.20.21 Transactions: GATE2016-1-22

https://gateoverflow.in/39644

Which one of the following is NOT a part of the ACID properties of database transactions?

A. Atomicity B. Consistency C. Isolation D. Deadlock-freedom

gate2016-1 databases transactions easy

3.20.22 Transactions: GATE2016-1-51

https://gateoverflow.in/39703

Consider the following two phase locking protocol. Suppose a transaction

T accesses (for read or write operations), a certain set of objects

$\{O_1, \ldots . O_k\}$. This is done in the following manner:

Step 1 . T acquires exclusive locks to $O_1, \ldots . , O_k$ in increasing order of their addresses.

Step 2 . The required operations are performed .

Step 3 . All locks are released

This protocol will

A. guarantee serializability and deadlock-freedom
B. guarantee neither serializability nor deadlock-freedom
C. guarantee serializability but not deadlock-freedom
D. guarantee deadlock-freedom but not serializability.

gate2016-1 databases transactions normal

3.20.23 Transactions: GATE2016-2-22

https://gateoverflow.in/39550

Suppose a database schedule S involves transactions $T_1, \ldots . . . . . , T_n$. Construct the precedence graph of S with vertices representing the transactions and edges representing the conflicts.If S is serializable, which one of the following orderings of the vertices of the precedence graph is guaranteed to yield a serial schedule?

A. Topological order
B. Depth-first order
C. Breadth- first order
D. Ascending order of the transaction indices

gate2016-2 databases transactions normal

3.20.24 Transactions: GATE2016-2-51

https://gateoverflow.in/39590

Consider the following database schedule with two transactions T_1 and T_2.

$S = r_2(X); r_1(X); r_2(Y); w_1(X); r_1(Y); w_2(X); a_1; a_2$

Where $r_i(Z)$ denotes a read operation by transaction T_i on a variable Z, $w_i(Z)$ denotes a write operation by T_i on a variable Z and a_i denotes an abort by transaction T_i.

Which one of the following statements about the above schedule is **TRUE**?

A. S is non-recoverable.
B. S is recoverable, but has a cascading abort.
C. S does not have a cascading abort.
D. S is strict.

gate2016-2 databases transactions normal

3.20.25 Transactions: GATE2019-11

https://gateoverflow.in/302837

Consider the following two statements about database transaction schedules:

I. Strict two-phase locking protocol generates conflict serializable schedules that are also recoverable.
II. Timestamp-ordering concurrency control protocol with Thomas' Write Rule can generate view serializable schedules that are not conflict serializable

Which of the above statements is/are TRUE?

A. I only B. II only C. Both I and II D. Neither I nor II

gate2019 databases transactions