Birla Institute of Technology & Science-Pilani
Hyderabad Campus
2nd Semester 2015-16
Database Systems (CS F212)
Comprehensive Examination (Regular)

Dt: 10.05.2016 AN Weightage: 40% Time: 3 Hrs. Type: Closed Book

PART-A

Note: (i) This part must be returned to the invigilator within 60 Mins. (**B**)
(ii) Mark your answers on this sheet only.
(iii) Overwriting not permitted.

Max Time: 60 Mins. Max. Marks=24X1=24 Total Marks obtained:

Name:
ID :

---

1. The number of Entity types participating in a relationship (in ER Modeling) is known as ____ of that relationship. [ ]
   (a) Entity set (b) cardinality (c) participation multiplicity (d) degree

2. A *weak entity type* cannot have more than one *partial* key. [True / False]

3. In *three-schema architecture*, at conceptual level there can be any number of conceptual schemas for a given database. [True /False]

4. Which of the following is true in case of *Tuple Relational Calculus*? [ ]
   (a) it is declarative (b) it is procedural (c) it is less expressive than SQL (d) it is commercial

5. When we use *generalization* for modeling Entity type hierarchies, which of the following is true?
   [ ]
   (a) some members of super type may not appear in any sub type
   (b) every member of super type will definitely appear in some sub type
   (c) it is always overlapping sub-grouping
   (d) a sub type cannot have specific relationship with other entity types

6. Which of the following operation is not a member of *complete set of Relational operations*?
   [ ]
   (a) selection (b) projection (c) cartesian product (d) union (e) intersection

7. The highest Normal Form satisfied by the relation R(A,B,C,D,E) with FDs
   { ABC--> DE; BC-->E;} is__ . [ ]
   (a) BCNF (b) 3NF (c) 2 NF (d) 1NF

8. If the relation R(A,B,C,D,E) with FDs {AB-->CD; D-->E} is decomposed into R1(A,B,C,D) and R2(B,E), then this decomposition is *lossless*. [True/False]

9. We can construct more than one secondary index on a relation. [True/False]

10. In the acronym *RAID* (in the context of disk storage of data), the letter '*A*' stands for _____.

11. In extendible hashing scheme, if we start with Global depth=3, and Local Depth=2, then how many buckets are allocated for storing data records, in the beginning? [ ]
(a) 4 (b) 8 (c) 2 (d) 1

12. For any functional dependency of the form X → Y, in a relation, BCNF requires that __ [ ]
a) X must be a key b) Y must be a key attribute
c) Y must be a key d) both X and Y must be key attributes

13. The time required to place the R/W head on the correct track from where we need to read the data, is known as__ [ ]
a) disk latency b) seek time c) track delay d) rotation delay

14. If the order (*p*) of the B+ tree is 7, then the minimum number of keys in any internal node is- [ ]
(a) 4 (b) 3 (c) 2 (d) 6

15. A *clustering index* is built on ___ attribute of a relation. [ ]
(a) ordering non-key (b) non-ordering key (c) non-ordering non-key (d) ordering key

16. The description that specifies the execution sequence of instructions in a set of transactions is called as a _____. [ ]
(a) Precedence pair (b) conflicting pair (c) schedule (d) system Log

17. Some conflict serializable schedules are not view serializable. [True/False]

18. Serializability can be implemented using__ [ ]
a) views b) indexes c) locks d) schedules

19. A transaction can never go to '***failed***' state from partially committed state. [True/False]

20. According to the SQL query optimization heuristics, a Cartesian product with subsequent selection condition (representing join condition) can be replaced by _____ [ ]
(a) project and division operation (b) join and projection (c) one join operation (d) one conjunctive select operation

21. In a Disk-pack(where data can be stored on both surfaces of disks) with uniform configuration, number of tracks per cylinder is same as number of surfaces in the Disk-pack. [True/False]

22. Assume that we have 7892 index blocks in the third level of multilevel Indexing structure. If the Block size is 512 Bytes and the blocking factor for Index file is 24, then we need _____number of index records at 4^{th} level. We assume unspanned records organization.
(fill-in the blank, do not write the formula here, just give the exact numerical value.).

23. W.r.t., storage capacity units, 10^{12} bytes is called as one *Tera byte*, and we call 10^{15} bytes as one _____ byte.

24. A secondary index built on a key field is always __ [ ]
(a) sparse (non-dense) (b) clustered (c) can be dense or sparse (d) dense

******

Birla Institute of Technology & Science-Pilani
Hyderabad Campus
2nd Semester 2015-16
Database Systems (CS F212)
Comprehensive Examination (Regular)

Dt: 10.05.2016 AN Weightage: 40% Time: 3 Hrs. Type: Closed Book

PART-B

Max Time: 3 hrs. Max. Marks=56

Note: (i) For this part write answers in Main answer booklet (ii) Answer the questions in the same order. (iii) sub-parts of a question must appear together. (iv) no additional sheets are provided hence use the space accordingly.

Q.1 Look at the following description related to Research and Sponsored Projects in an Educational Institution.

Institute has some sponsored Projects. Projects will have one Faculty as Principal Investigator (PI), and can have one or more faculty member as Co-PIs. Some projects will not have any Co-PI. A Project has ProjectID (unique), Project name, Budget, and Duration as attributes. Faculty are identified by FacultyID (unique) and have Name, Dept , Designation as other attributes. A Project is funded by only one Funding agency (like UGC, DIT, DST etc.) which has- Agency name (unique), Head, Location (with street, city, and state as sub components). A faculty, as a PI can have zero to any number of projects. Similarly a faculty as a Co-PI can have zero to any number of projects. A funding agency might have funded one or more projects.

(i) First draw an ER diagram for the above requirement. Assume necessary data which is missing in the question. The model should include- Entity types, relationships, min-max, cardinality, participation, and other relevant constraints.
(ii) Then design relational schemas to capture the data represented in the ER diagram you have drawn. [4 + 4 = 8]

Q.2 Look at the following Database schema.

River(<u>rid</u>, rname, length, type)
State(<u>sname</u>, capital, population)
RiverState(<u>rid</u>, <u>sname</u>) //This relation captures info about what river flows through what states, and here *rid* is FK to *rid* of *River*, and *sname* is FK to *sname* of *State*

Now, write both Relational Algebra and SQL query expressions to:

(i) Get the *rid* and *rname* for those rivers which are perennial type and do not flow through any state with population greater than 4 crores. [2 + 2 =4]
(ii) Get the *stname* and *capital* for those states having population greater than the average population of all states and have at least 3 rivers flowing through. [2 + 2= 4]

Note: You don't have to rename the attributes in the result, use only DML statements, and don't use outer joins. Do not define new tables or views.

Q.3. (a) For the following schemas (i) identify the keys (ii) give the highest NF satisfied (with reasons), then, (iii) if it is not in BCNF give decomposition to bring it to BCNF. [2X2=4]

(i) R1(A,B,C,D) { A→B; B→C; C→D}
(ii) R2(A,B,C,D,E) { AB→CDE; DE→ABC; E→C}

(b) In an educational institution, we store data about **Courses** and **Faculty**. The entity **Course** has *cid* (unique), *cname*, and *credits* as attributes and **Faculty** has *facid* (unique), *name, address, designation* and *phone* as attributes. The *cid* can determine *cname* and *credits*; further *facid* can determine all attributes of Faculty entity type. Each course must be taught by one Faculty and a Faculty can teach many courses. Some faculty are not allotted to any course. Now if we design a single table ***Course_Faculty*** (cid, cname, credits, facid, name, designation, address, phone) to capture all the data, with cid as the PK, (since cid→facid), brief on the problems we face with this design and also explain how *insert* and *delete* anomaly occur in this situation. [4]

Q.4 (a) Determine if the following schedule is (conflict) serializable, for the concurrent transactions 1, 2 , 3and 4. The data items are A, B, and C. Draw the *precedence graph*. Give your comments to justify the answer.

Schedule 1: {$r_2(A)$; $r_3(B)$; $r_2(C)$; $r_4(A)$; $w_4(A)$; $w_2(C)$; $w_3(B)$; $r_1(A)$; $r_4(C)$; $r_2(B)$; $r_1(B)$; $w_1(A)$;}

Note: Here, $r_1(A)$; - means that the transaction-1 reads data item A

$w_2(B)$; - means that the transaction-2 writes data item B [5]

(b) Give a Concurrent schedule with three transactions which is not *conflict serializable*, but *view serializab*le. [3]

Q.5 Assume that we use ***Extendible hashing*** technique in some situation and we use the hash function (K mod 8). Assume that a bucket (one block) can accommodate 2 records. Now insert the records with following keys in same order, and show the dynamic structure of the hashing scheme after each insertion with simple diagrams. Start with Global depth as 3, and local depth as 2. Keys to be inserted are: 7, 10, 17, 32, 44, 35, 80, 63.

(Note: never go for overflow buckets. Always resolve collision by increasing the local depth. Complete working is required) [6]

(b) Brief on how time-stamp based ***wound-wait*** deadlock prevention scheme works. [2]

Q.6 (a) Assume that we have two relations- EMP(eid, ename, sal, dno, emploc) and DEPT(dnum, dname, dloc). Further, *dno* in EMP is KF to *dnum* in DEPT. Now explain the difference between the following strategies for joining these two tables based on emploc=dloc.

Strategy 1: If there exist no indexes on EMP and DEPT.

Strategy 2: If there exists an index on *dloc* of DEPT table. [5]

(give simple suitable *pseudo code* for the joining algorithm for the above two cases)

(b) How do we recover from a failure in ***shadow paging*** based DB recover scheme . [3]

Q.7 (a) For the following SQL query, give two different ***query trees*** (need not be optimized).

```
SELECT   S.sid, S. name, C.cname
FROM   Student as S, Company as C
WHERE   C.cid=S.cid and S.branch='CIVIL' and C.loc='Pune';
```

[5]

(b) For what sort of query requirements, *Bit-map indices* are suitable and why? [3]

*****