Reg.No. \_\_\_\_\_\_\_\_\_\_\_\_



**UNIVERSITY**

(Karunya Institute of Technology & Sciences)

(Declared as Deemed-to-be University under Sec.3 of the UGC Act, 1956)

**End Semester Examination – Nov/Dec – 2016**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Semester :** | **2016-17 ODD** |
| **Code :** | **14CS3005** | **Duration :** | **3hrs** |
| **Sub. Name :** | **ADVANCED DATABASE SYSTEMS** | **Max. marks :** | **100** |

**ANSWER ALL QUESTIONS (5 x 20 = 100 Marks)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Q. No** | **Sub Div.** | **Questions** | **Course**  **Outcome** | **Marks** |
| 1. | a. | Draw the relational algebra tree for the following query. List the different approaches available for the Join operation and calculate the corresponding I/O cost.  Select E.Empname, E.ManagerId, D.Deptname  From Emp E, Dept D  Where E.did = D.did  Use the following statistics for I/O cost calculation:   * + Emp: Each tuple is 80 bytes long, 40 tuples per page, 300 pages.   + Dept: Each tuple is 50 bytes long, 90 tuples per page, 100 pages.   Note: Write down all the assumptions made in calculating the cost. | CO1 | **10** |
| b. | Give specific examples to illustrate the impact of ‘pushing selections down’ in a relational algrbra tree. | CO1 | **5** |
| c. | Discuss the following evaluation types : Materialized and Pipelined Evaluation. Which type is more beneficial and why? | CO1 | **5** |
| **(OR)** | | | | |
| 2. | a. | Consider the following schedule S1. Draw the precedence graph for S1 and state whether the schedule is conflict serializable or not. Show all the intermediate steps of converting the given schedule to a serial schedule.  S1: r1(X); r2(Z); r3(X); r1(Z); r2(Y); r3(Y); w1(X); w2(Z); w3(Y); w2(Y) | CO2 | **8** |
| b. | When do two actions on the same data object conflict? Explain the different types of anomalies with suitable examples. | CO2 | **8** |
| c. | What is 2-phase locking protocol? Where is it used? | CO3 | **4** |
| 3. | a. | Use lock-based concurrency control mechanism and verify whether the following schedule is valid. Apply the relevant locks and explain how the concurrency control manager will handle the schedule.  S : r1(A); r2(B); r3(C); r1(B); r2(C); r3(A); w1(A); w2(B); w3(C); commit1; commit2; commit3 | CO3 | **8** |
|  | b. | Insert locks and draw wait for graph for the schedule given in question 3a and state whether it is deadlock-free schedule. | CO3 | **6** |
|  | c. | Describe the specialilzed locking technique for B+ tree index structure with suitable example. | CO3 | **6** |
| **(OR)** | | | | |
| 4. | a. | What is the purpose of ‘log’ information in recovery process? What are the different types of LSN records created in ARIES algorithm? | CO2 | **5** |
|  | b. | Explain the steps during the Analysis, Redo and Undo phases of ARIES algorithm using the following log information.   |  |  | | --- | --- | | **LSN** | **LOG** | | 00 | begin\_checkpoint | | 10 | end\_checkpoint | | 20 | Update: T1 writes P5 | | 30 | Update: T3 writes P2 | | 40 | Update: T1 writes P4 | | 50 | T1 abort | | 60 | Update: T3 writes P4 | | 70 | T3Commit | | 80 | T3 end | | 90 | Update : T2 writes P2 | | 100 | T2 abort | | **X** Crash, Restart | | | CO2 | **15** |
| 5. | a. | Consider the following relations and query :  Emp(Eid, Ename, Designation, Managerid, Salary, Dno)  Dept(Dno, Dname, Mgrname, location)  Select E.ename, D.mgrname from Emp E, Dept D where D.dname=‘Finance’ AND E.Dno=D.dno  Suggest suitable indexes to execute the query in an efficient manner (Provide 2 alternatives) | CO1 | **6** |
|  | b. | Automatic index selection is a hard problem. Why? | CO1 | **4** |
|  | c. | Why is it necessary to obtain certificates from certification authorities like Verisign? Explain the process of obtaining the certificate. | CO4 | **10** |
| **(OR)** | | | | |
| 6. | a. | What are objects, subjects, security classes and clearances in Mandatory access control? Discuss the properties of Bell-LaPadula. | CO2 | **6** |
|  | b. | What shoud be the class security level of the user to read/update each row of the following table? What do you mean by ployinstantiation and multilevel table? Explain using suitable examples. (Use Bell-LaPadula model).   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Studid** | **Sname** | **Dept** | **CGPA** | **Security Class** | | 1 | Sam | CSE | 7.5 | TS | | 2 | Veenu | IT | 8.9 | C | | 3 | Joseph | ECE | 9.2 | S | | 4 | Ruth | EIE | 9.5 | C | | CO2 | **8** |
|  | c. | What is an authorization graph? Explain SQL’s GRANT and REVOKE commands. | CO2 | **6** |
| 7. | a. | Why data partitioning is important? Explain the different data partitioning technique with suitable illustration. | CO4 | **6** |
|  | b. | Consider the following query :  Select E.Empname, E.ManagerId, D.Deptname  From Emp E, Dept D Where E.did = D.did  Employee table is available at London and Department table at Newyork. The resultant table is required at London. What are the different approaches available to join these 2 tables? Calculate the I/O cost by using the statistics given in Question1a. | CO4 | **10** |
|  | c. | What is the difference between synchronous and asynchronous replication? Why asynchronous replication has gained popularity? | CO4 | **4** |
| **(OR)** | | | | |
| 8. | a. | Write about the two index structures available for text retrieval. | CO4 | **10** |
|  | b. | What are the ways to measure the document similarity? (i) Use the vector space model and calculate IDF for each document listed in the following table (ii) determine the weight vector of each document. (iii) Create an inverted index and evaluate the Boolean queries for (i) Pencil AND Pen AND Rubber, (ii) Books AND (Note OR Rubber).   |  |  | | --- | --- | | **Doc ID** | **Terms** | | 1 | Books, Pen, Rubber | | 2 | Pencil, Note, color pen, Sharpner | | 3 | Pen, Pencil, Box, Bag | | 4 | Bag, Scale, Clips, Rubber | | CO4 | **6** |
|  | c. | What is the vector space model and what are its advantages? | CO4 | **4** |
|  | | **Compulsory:** |  |  |
| 9. | a. | What are the characteristics of spatial data? What are the differences between spatial range, nearest neighbor and spatial join queries. | CO4 | **5** |
|  | b. | What is space-filling curve and how can it be used to design a spatial index? Describe spatial index structures based on space filling curves. | CO4 | **10** |
|  | c. | Consider the following points and regions, with x and y coordinates each of 3 bits. A: (3,2), B:(4,2), C(2,3), D: rectangle corners (0,6), (1,6), (1,7), (0,7), E: rectangle corners (5,3), (6,3), (6,5), (5,5). Use Z-ordering space-filling curve to represent these points and regions. | CO4 | **5** |

ALL THE BEST