Reg. No. : | | | | | | | | | | | | |

Question Paper Code : 70379

B.E./B.Tech. DEGREE EXAMINATION, NOVEMBER/DECEMBER 2021.

Third/Fifth/Eighth Semester

Computer Science and Engineering

CS 6302 – DATABASE MANAGEMENT SYSTEMS

(Common to Information Technology/Mechanical and Automation Engineering)

(Regulations 2013)

Time : Three hours Maximum : 100 marks

Answer ALL questions.

PART A — ($10 \times 2 = 20$ marks)

1. What are the disadvantages of file processing system?
2. Explain entity relationship model.
3. What is a query execution plan?
4. Which cost component are used most often as the basis for cost function?
5. What is meant by concurrency control?
6. Give an example of Two phase commit protocol.
7. List out the mechanisms to avoid collision during hashing.
8. What are the disadvantages of B Tree over B+ Tree?
9. Define Threats and risks.
10. What is Association rule mining?

PART B — ($5 \times 13 = 65$ marks)

11. (a) (i) With help of a neat block diagram explain the basic architecture of a database management system. (6)

(ii) What are the advantages of having a centralized control of data? Illustrate your answer with suitable example. (7)

Or

(b) A Car rental company maintains a database for all vehicles in its current fleet. For all vehicles, it includes the vehicle identification number license number, manufacturer, model, date of purchase and color. Special data are included for certain types of vehicles. (13)

Trucks: Cargo capacity

Sports Cars: horsepower, renter age requirement

Vans: number of passengers

Off-road vehicles, ground clearance, drivetrain (four-or two-wheel drive)

Construct an ER model for the car rental company database.

12. (a) Consider a student registration database comprising of the below given table schema. (13)

Student File

Student Number	Student Name	Address	Telephone

Course File

Course Number	Description	Hours	Professor Number

Professor File

Professor Number	Name	Office

Registration File

Student	Course	Date
Number	Number	

Consider a suitable sample of tuples/records for the above mentioned tables and write DML statements (SQL) to answer for the queries listed below.

(i) Which courses does a specific professor teach?

(ii) What courses are taught by two specific professors?

(iii) Who teaches a specific course and where is his/her office?

(iv) For a specific student number, in which courses is the student registered and what is his/her name?

(v) Who are the professors for a specific student?

(vi) Who are the students registered in a specific course?

Or

(b) Discuss about the Join order optimization and Heuristic optimization algorithms. (13)

13. (a) Discuss the violations caused by each of the following: dirty read, non repeatable read and phantoms with suitable example. (13)

Or

(b) Explain why timestamp-based concurrency control allows schedules that are not recoverable. Describe how it can be modified through buffering to disallow such schedules. (13)

14. (a) Briefly explain RAID and RAID levels. (13)

Or

(b) Briefly explain about B+ tree index file with example. (13)

15. (a) Neatly write the K-means algorithm and show the intermediate results in clustering the below given points into Two clusters using K-means algorithm.

P1 (0,0), P2: (1,10), P3: (2,20), P4: (1,15), P5: (1000, 2000), P6 : (1500, 1500), P7 : (1000,1250). (13)

Or

(b) Discuss about the Access control mechanisms and Cryptography methods to secure the Databases. (13)

PART C — ($1 \times 15 = 15$ marks)

16. (a) Consider the relation schema given in Figure 1. Design and draw an ER diagram that capture the information of this schema. (5)

Empolyee (empo, name, office, age)
Books (isbn, title, authors, publisher)
Loan (empno, isbn, date)

Fig.1

Write the following queries in relational algebra and SQL.

(i) Find the names of employees who have borrowed a book published by McGraw-Hill. (5)

(ii) Find the names of employees who have borrowed all books published by McGraw-Hill. (5)

Or

(b) Trace the result of using the Apriori algorithm on the grocery store example with support threshold $s = 33.34\%$ and confidence threshold $c = 60\%$. Show the candidate and frequent itemsets for each database scan. Enumerate all the final frequent itemsets. Also indicate the association rules that are generated and highlight the strong ones, sort them by confidence. (15)

Transaction ID	Items
T1	HotDogs, Buns, Ketchup
T2	HotDogs, Buns
T3	HotDogs, Coke, Chips`
T4	Chips, Coke
T5	Chips, Ketchup
T6	HotDogs, Coke, Chips