**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI**
**(END SEMESTER EXAMINATION)**

**CLASS: IMSc** **SEMESTER : VI**
**BRANCH: MATHEMATICS** **SESSION: SP/2023**

**SUBJECT: CS325 DATABASE MODELLING**

**TIME: 3 Hours** **FULL MARKS: 50**

**INSTRUCTIONS:**
**1. The question paper contains 5 questions each of 10 marks and total 50 marks.**
**2. Attempt all questions.**
**3. The missing data, if any, may be assumed suitably.**
***4*. Before attempting the question paper, be sure that you have got the correct question paper.**
**5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.**

---

| | | | CO | BL |
|---|---|---|---|---|
| Q.1(a) | State advantages of using database over file system for storing data. How DDL differs from DML? Write the ER semantic tools required for view integration of database design. | [2+2+1] | 1 | 2 |
| Q.1(b) | Consider the following scenario and draw an E-R diagram for Robert.<br>Robert has decided to store information about musicians who perform on its albums (as well as other company data) in a database. Each album that is recorded on the Robert label has a title, a copyright date, a format (e.g., CD or MC), and an album identifier. Each album has a number of songs on it, but no song may appear on more than one album. Each album has exactly one musician who acts as its producer. A musician may produce several albums, of course. Each musician that records at Robert has a SerialNo, a name, an address, and a phone number. Each address is made up of a street, a city and a postcode. Each musician may play several instruments, and a given instrument may be played by several musicians. Each instrument that is used in songs recorded at Robert has an instrumentID, name (e.g., guitar, piano, flute) and a musical key (e.g., C, B-flat, E-flat). Each song is performed by one or more musicians, and a musician may perform a number of songs. Each song recorded at Robert has a title and an author | [5] | 2 | 5 |
| Q.2(a) | Define entity integrity and referential integrity with suitable examples. Explain different types of conflicts arisen from schema diversity. | [2+3] | 2 | 2 |
| Q.2(b) | Draw a class diagram showing the structure of data about employees of a given company. The employees attributes include name, street address, city, state, zip, and an id number. The employees also have an annual salary. Departments have names and are composed of a collection of employees, but each employee can be in only one department. Employees work on one or more projects, which also have names. Projects may have multiple employees assigned to them. Include multiplicity, and attributes in your diagram. | [5] | 3 | 4 |
| Q.3(a) | The relation scheme Student_Performance (name, courseNo, rollNo, grade) has the following functional dependencies:<br>name, courseNo → grade<br>rollNo, courseNo → grade<br>name → rollNo<br>rollNo → name<br>Determine the highest normal form of the relation scheme with proper explanation. | [3] | 3 | 5 |
| Q.3(b) | A table has fields F1, F2, F3, F4 and F5, with the following functional dependencies: F1 → F3, F2 → F4, (F1, F2) → F5. Determine the highest normal form of the table with proper explanation. | [3] | 3 | 5 |
| Q.3(c) | Explain the anomalies that can occur when a database is not normalized, with suitable examples. | [4] | 3 | 2 |

**PTO**

| | | | | |
|---|---|---|---|---|
| Q.4(a) | Define data Warehousing and data mining. Write the core requirements of data warehousing. | [2+3] | 4 | 2 |
| Q.4(b) | Compare the following.<br>i. OLTP vs OLAP<br>ii. Apriori Algorithm vs FP growth | [5] | 4 | 3 |
| Q.5(a) | Apply Apriori algorithm on a book store example with support threshold, s=33.34% and confidence threshold, c=60%, where H, B, K, C and P are different books purchased by customers. | [5] | 4 | 5 |

| Transaction Id | Items |
|---|---|
| T1 | H,B,K |
| T2 | H, B |
| T3 | H, C, P |
| T4 | P, C |
| T5 | P, K |
| T6 | H, C, P |

Find all frequent itemsets by using Apriori Algorithm. Show all the association rules that are generated. Write all the strong association rules

| | | | | |
|---|---|---|---|---|
| Q.5(b) | Write importance of CASE tools. Why is XML known as semi-structured format? Write the features of XML. | [2+1+2] | 5 | 2 |

**::::::26/04/2023::::::M**