Course Title: Database Management System **Full Marks: 60+20+20**
Course No: BIT202 **Pass Marks: 24+8+8**
Nature of the Course: Theory + Lab **Credit Hours: 3**
Year/Semester: Second/Third

Course Description: The course covers the basic concepts of databases, database system concepts and architecture, data modeling using ER diagram, relational model, SQL, relational algebra and calculus, normalization, transaction processing, concurrency control, and database recovery.

Course Objectives: The main objective of this course is to introduce the basic concepts of database, data modeling techniques using entity relationship diagram, relational algebra and calculus, basic and advanced features SQL, normalization, transaction processing, concurrency control, and recovery techniques.

Detail Syllabus

Chapters / Units / Content Depth	Teaching Hours
Unit 1: Database Concepts and Architecture (4 Hours) Database *(Introduction, Examples)*, Database Management System *(Introduction, Advantages, Examples)*; Database Users(*Actors on the scene: Database Administrators, Database Designers, End Users(Causal, Naïve, Sophisticated, Standlone), System Analyst and Standalone Programmers, Workers behind the scene: DBMS System Designers and Implementers, Tool Developers, Operators and maintenance personnel)* and Benefits of Databases;	**1 Hour**
Data Models *(Hierarchical, Network, ER, Relational, Object):* , Schemas, and Instances; Three-Schema Architecture *(Internal, Conceptual and External View)* and Data Independence *(Logical and Physical)*; Database Languages *(DDL, DML, SDL, VDL)* and Interfaces *(Menu Based, Mobile Apps, Form Based, GUI, Natural Language, Keyword Based, Search Input Output, Interfaces for Parametric users and DBA);* The Database System Environment*(DBMS Component Modules, Database System Utilities, Tools, Application Environments, and Communications Facilities);*	**2 Hours**
Centralized and Client/Server Architectures for DBMSs *(One-tier, Two-tier, Three-tier, N-tier);* Classification of Database Management Systems *(Single and Mutli user, Centralized and Distributed, Homogeneous and Heterogeneous, General Purpose and Special Purpose);*	**1 Hour**

Unit 2: Data Modeling Using the Entity-Relational Model (5 Hours) Using High-Level Conceptual Data Models for Database Design *(Requirement Analysis, Conceptual Design, Logical Design, Physical Design)*;	**0.5 Hour**
Entity Types, Entity Sets, Attributes *(Composite Vs. Simple, Single Valued Vs. Multivalued, Stored Vs. Derived)* and Keys; Relationship Types, Relationship Sets, *Relationship Instances*, *Relationship Degree*, Roles, and Structural Constraints*(Cardinality Ratio, Participation Constraint)*; Weak Entity Types *(Weak Entity and Partial Keys);* ER Diagrams, Naming Conventions*(Attribute, Entity, Relationship)*, and Design Issues; Relationship Types of Degree Higher Than Two; *ER-to-Relational Mapping*	**3.5 Hours**
Enhanced Entity–Relationship (EER): Subclasses, Superclasses, and Inheritance; Specialization and Generalization; Constraints (*Overlapping, Disjoint*) and Characteristics of Specialization and Generalization;	**1 Hour**
Unit 3: The Relational Data Model and Relational Database Constraints (5 Hours) Relational Model Concepts *(Domains, Attributes, Tuples, and Relations; Characteristics of Relation)*; Relational Model Constraints *(Domain Constraints, Key Constraints, Constraints on Null Values)* and Relational Database Schemas *(Relational Database state)*; *(Entity Integrity, Referential Integrity, and Foreign Keys);* Update Operations, Transactions, and Dealing with Constraint Violations *(Inset, Delete Update Operations) (Restrict/Cascade/set null/ set default)*;	**2.5 Hours**
Basic Relational Algebra Operations *(Unary: Select, Project, Rename; Binary: Binary: Set Theory(Union, Intersection, Set Difference), Cartesian Product, Join (Natural/Theta), Outer Join (Left/Right/Full);*	**2.5 Hours**

Unit 4: SQL (10 Hours) Data Definition and Data Types *(Attribute Data Types and Domains Create Database, Create Table, Drop Table, Drop Constraint, Drop Database);* Specifying Constraints *(Attribute Constraints, Attribute Defaults, Key and Referential Integrity Constraints, Naming Constraints, Constraint using CHECK);*	**2 Hours**
Basic Retrieval Queries *(SELECT-FROM-WHERE Structure, Ambiguous Attribute Names, Aliasing, Renaming, Tuple Variables; Unspecified WHERE clause, Use of * in Select, Substring Pattern Matching using LIKE, Arithmetic Operators, Order by Clause);*	**3 Hours**
Complex Retrieval Queries*; (Where Clause using IS NULL, Logical Connectives, Nested Query, Correlated Nested Query, Using BETWEEN, IN and EXISTS Clauses in Where, Renaming Attributes, JOIN, Natural JOIN, OUTER JOIN (Left/Right), Aggregate Functions, GROUP BY and HAVING Clause);*	**3 Hours**
INSERT, DELETE, and UPDATE Statements;	**1 Hour**
Views *(CREATE, DROP);*	**1 Hour**
Unit 5: Relational Database Design (7 Hours) Relational Database Design Using ER-to-Relational Mapping *(Mapping of Regular Entity, Weak Entity, Relationship Types, Multivalued Attributes, N-ary Relationship);*	**1 Hour**
Informal Design Guidelines for Relational Schemas *(Semantics of Attributes in Relations, Redundant Information in Tuples and Update Anomalies, NULL Values in Tuples, Generation of Spurious Tuples);*	**1 Hour**
Functional Dependencies *(Definition, Inference Rules for Functional Dependencies, Armstong's Axioms, Attribute Closure, Equivalence of Functional Dependencies, Minimal Sets of Functional Dependencies);* Normal Forms Based on Primary Keys *(First Normal Form, Second Normal Form, Third Normal Form);* General Definitions of Second and Third Normal Forms;	**2.5 Hours**
Boyce-Codd Normal Form; Multivalued Dependency and Fourth Normal Form;	**2 Hours**
Properties of Relational Decomposition *(Attribute Preservation, Dependency Preservation, Loss-Less/ Non-Additive Join)*	**0.5 Hour**

Unit 6: Transaction Processing and Concurrency Control, and Recovery (8 Hours) Introduction to Transaction Processing; *(Single User vs. Multi User, Read/ Write Operations, Need for Concurrency Control: Lost Update Problem, Temporary Update (or Dirty Read) Problem, Incorrect Summary Problem, Unrepeatable Read Problem, Need for Recovery);*	**2 Hours**
Transaction and System Concepts *(Transaction States, System Log, Commit Point of Transaction)*; Desirable Properties of Transactions (*Atomicity, Consistency, Isolation, Durability)*;	**2 Hours**
Schedules, Conflicting Operations in Schedule, Characterizing Schedules based on Recoverability, Characterizing Schedules based on Seriablizility, Serial, Nonserial, Conflict Serializable Schedules, Testing for serializability of Schedule, Using Serializability for Concurrency Control;	**2 Hours**
Concurrency Control Techniques; Two-Phase Locking *(Types of Lock, Basic, Conservative, Strict, and Rigorous Two-Phase Locking, Deadlock and Starvation, Deadlock Prevention, Deadlock Detection)* and Timestamp Ordering *(Timestamp, Read Timestamp, Write Timestamp, Basic Timestamp Ordering, Strict Timestamp Ordering);*	**2 Hours**
Unit 7: Database Recovery Techniques (3Hours) Recovery Concepts *(Recovery outline and categorization of recovery algorithms; Caching (Buffering) of disk blocks; Write-ahead logging, Steal/no-steal, and Force/no-force; Checkpoints and Fuzzy Checkpointing; Transaction rollback and cascading rollback);*	**1Hour**
NO-UNDO/REDO Recovery Based on Deferred Update; Recovery Technique Based on Immediate Update; Shadow Paging; Database Backup and Recovery from Catastrophic Failures;	**2Hours**
Unit 8: NoSQL (3 Hours) Structured and Unstructured Data; Introduction to NoSQL Databases *(NoSQL Database, Types of NoSQL Database, Advantage of NoSQL);*	**1 Hour**
Discussion of basic architecture of Hbase, Cassandra and MongoDb;	**2 Hours**

Text Books:

1. **Fundamentals of Database Systems; Seventh Edition; Ramez Elmasri, Shamkant B. Navathe; Pearson Education**
2. Database System Concepts; Sixth Edition; AviSilberschatz, Henry F Korth, S Sudarshan; McGraw-Hill
3. NoSQL for Dummies; Adam Fowler; John Wiley & Sons, Inc.

Reference Books:

1. Database Management Systems; Third Edition; Raghu Ramakrishnan, Johannes Gehrke; McGraw-Hill
2. A First Course in Database Systems; Jaffrey D. Ullman, Jennifer Widom; Third Edition; Pearson Education Limited

Laboratory Work Manual

The laboratory work includes design and implementation of database. Students should practice logical design of databases (ER Models) using case tools like draw.io, MS-Visio etc. The so designed database should be implemented by writing database programs using MYSQL, ORACLE etc. The database programs should include the DDL and DML statements like;

- Create, Drop, Alter, Rename of database/table/constraints
- INSERT, DELETE, and UPDATE Statements
- Specifying constraints (primary key, foreign key, referential integrity etc.)
- Basic and complex DML queries using select from where clause including the use of Order By, LIKE, IN, Exists, Between, Logical Connectives, LIMIT etc.
- Join Operations
- Set Theory Operations Union, Intersect, Set Difference
- Nested Queries
- Aggregate functions
- Group By Having Clause
- Create Views / Drop Views
- Concepts of transaction, Backup and recovery

In addition, the students should practice and get familiar with the basics of NoSQL Database.

Model Question
Tribhuvan University
Institute of Science and Technology

Course Title: Database Management System **Full Marks:** 60
Course No: BIT202 **Pass Marks:** 24
Level: BIT Second Year/ Third Semester **Time: 3 Hrs**

Section A
Long Answer Questions

*Attempt any **TWO** questions.* *[2×10=20]*

1. Consider a database system with following schemas; [10]

Hospital(hname, haddress, hspecilaity)
Doctor(did, dname, dspecilization,)
Worksat(did, hname, workinghrs)
Pharmacy(phname, hname, no_of_sales, total_revenue)

Now write SQL statements and relational algebra statements for following queries;
   a. Select name of all doctors having specialization "gyno".
   b. Select the name and address of hospital where working hours is "day".
   c. Using natural join select the name of doctors whose working hours are "night".
   d. Find the average salary of the doctors.
   e. Find names of hospital and their pharmacy which generate revenue more than 10000. Sort the result in descending order on the basis of pharmacy name.
2. What do you mean by weak entity and partial key? Design an ER diagram for following scenario; [3+7]

In film industry, producers produce movies. Producers have their name, age and budget as attributes. They are uniquely identified by prod_id. All the movies have their title, year, and release date. No movies can have same title. Every movies must be played by actor. An actor can play many movies. Actors have Fname and Lname to uniquely identify them. The actors have charge_rate as well. A single movie can have many producers and a producer can produce many movies.

3. Define serial, non-serial and serilizable schedules with example. How can you test serializability in a schedule, illustrate with an example? [6+4]

Section B
Short Answer Questions

*Attempt any **EIGHT** questions.* *[8×5=40]*

4. What is data independence? How three schema architecture ensures logical and physical data independence? [1+4]
5. What do you mean by generalization specialization in EER? Illustrate with examples. [5]
6. Define domain constraint and referential integrity constraint. [5]
7. Why normalization is needed in database design? Describe third normal form with an example. [1.5+3.5]
8. Show how lost update and incorrect summary problem might occur in concurrent execution of transactions. [5]

9. Define Outer Join in SQL. Given following relations, show the results of left and right out Joins. [1.5+3.5]

Employee

Eid	Ename	Address	Dno
1	Ram	KTM	111
2	Rita	PKR	222
3	Hari	KTM	333

Department

Dno	Dname
111	HRM
222	Admin
444	Account

10. What is shadow paging? How it is used for database recovery? [3+2]
11. Differentiate structured and unstructured data. List the advantages of NoSQL. [2.5+2.5]
12. During ER-to-Relational Mapping, show how you map 1:N and N:M relationship into a relation? [5]