CSPC2004 DATABASE ENGINEERING (3-0-0)

Course Learning Objectives
CLO 1. To provide a strong foundation in database concepts, technology, and practice.
CLO 2. To practice SQL programming through a variety of database problems.
CLO 3. To Understand the relational database design principles.
CLO 4. To demonstrate the use of concurrency and transactions in database
CLO 5. To design and build database applications for real world problems.
CLO 6. To become familiar with database storage structures and access techniques.

Module-1 (8 hours)
Introduction to Databases and overview of Database Languages and architecture:
Overview, Database System vs File System, Overall Database Structure, Characteristics of database approach, Advantages of using the DBMS approach, History of database applications, Database System Concept, Types of users, Database design lifecycle, Data Model Schema and Instances, Types of data model, Three schema architecture, Data Independence and Database Language and Interfaces, Data Definitions Language, DML, the Database System environment,
Data Modeling Using the Entity Relationship Model:
ER Model Concepts, Notation for ER Diagram, Mapping Constraints, Keys, Concepts of Super Key, Candidate Key, Primary Key, Specialization, Generalization, Aggregation, Reduction of an ER Diagrams to Tables, Extended ER Model, Relationship of Higher Degree.

Module-2 (8 hours)
Introduction to Relational Data Model:
Relational Model Concepts, Codd's rules, Integrity Constraints, Entity Integrity, Referential Integrity, Key Constraints, Domain Constraints and relational database schemas, Update operations, transactions, and dealing with constraint violations.
Relational Algebra and Calculus:
Selection and projection, set operations, renaming, Joins, Division, syntax, semantics. Operators, grouping and ungrouping, relational comparison. Calculus: Tuple relational calculus, Domain relational Calculus, calculus vs algebra, computational capabilities, Relational Database Design using ER-to- Relational mapping.
Structured Query Language:
Introduction to SQL- Characteristics, advantages of SQL, Types of SQL commands, SQL data definition and data types, specifying constraints in SQL, basic retrieval queries in SQL, INSERT, DELETE, and UPDATE statements in SQL, Additional features of SQL. More complex SQL retrieval queries, Specifying constraints as assertions and action triggers, Views and Indexes, Queries and Sub Queries. Aggregate Functions. Insert, Update and Delete Operations, Joins, Unions, Intersection, Minus, Cursors, Triggers, Procedures in SQL/PL SQL.

Module-3 (9 hours)
Normalization:
Database Design Theory – Introduction to Normalization using Functional and Multivalued Dependencies: Informal design guidelines for relation schema, Functional Dependencies, Normal Forms based on Primary Keys, Second and Third Normal Forms, Boyce-Codd

Normal Form, Multivalued Dependency and Fourth Normal Form, Join Dependencies and Fifth Normal Form. Examples on normal forms.

Normalization Algorithms:
Inference Rules, Equivalence, and Minimal Cover, Properties of Relational Decompositions, Algorithms for Relational Database Schema Design, Nulls, Dangling tuples, and alternate Relational Designs, Further discussion of Multivalued dependencies and 4NF, Other dependencies and Normal Forms.

Module-4 (8 hours)
File Structures, Hashing and Indexing:
Introduction, placing file records on disk, hashing techniques, Parallelizing Disk Access Using RAID Technology, indexing structures for files- Types of Single-Level Ordered Indexes, Dynamic Multilevel Indexes Using B-Trees and B+-Trees, Indexes on Multiple Keys
Query Processing, optimization and Database tuning:
Translating SQL Queries into Relational Algebra, Algorithms for External Sorting, Algorithms for SELECT and JOIN Operations, Algorithms for PROJECT and Set Operations, Implementing Aggregate Operations and OUTER JOINs, Combining Operations Using Pipelining, Using Heuristics in Query Optimization, Using Selectivity and Cost Estimates in Query Optimization, Physical Database Design in Relational Databases, an Overview of Database Tuning in Relational Systems

Module-5 (7 hours)
Transaction Processing:
Introduction to Transaction Processing, Transaction and System concepts, ACID properties, Testing of Serializability, Serializability of Schedules, Conflict & View Serializable Schedule, Recoverability, Recovery from Transaction Failures, Log Based Recovery, Checkpoints, Deadlock Handling
Concurrency Control Techniques:
Locking Techniques for Concurrency Control, Time Stamping Protocols for Concurrency Control, Validation Based Protocol, Multiple Granularity, Multi Version Schemes, Recovery with Concurrent Transaction.

Course Outcomes
At the end of the course the student will be able to:
CO1 Identify, analyze and define database objects, enforce integrity constraints on a database using RDBMS.
CO2 Use Structured Query Language (SQL) for database manipulation and also demonstrate the basic of query evaluation.
CO3 Design and build simple database systems and relate the concept of transaction, concurrency control and recovery in database
CO4 Develop application to interact with databases, relational algebra expression. CO 5. Develop applications using tuple and domain relation expression from queries

Textbooks
1. Fundamentals of Database Systems, Ramez Elmasri and Shamkant B. Navathe, 7th Edition, 2017, Pearson.
2. Database management systems, Ramakrishnan, and Gehrke, 3rd Edition, 2014, McGraw Hill

3. C. J. Date, A. Kannan and S. Swamynathan, "An Introduction to Database Systems", 8th ed, Pearson Education, 2006

Reference Books:
Abraham Silberschatz, Henry F. Korth and S. Sudarshan's Database System Concepts 6th Edition Tata Mcgraw Hill Education Private Limited.

Weblinks and Video Lectures (e-Resources):
https://www.youtube.com/watch?v=ywTn9qHyI9I
https://www.youtube.com/watch?v=_qbKMdqQS6E
https://www.youtube.com/watch?v=PqPkYmRSQ_w